A More Accurate Measurement Model for Fault-Tolerant Quantum Computing

by

Yingkai Ouyang

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2008

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Aliferis, Gottesman and Preskill [1, 2] reduce a non-Markovian noise model to a local noise model, under assumptions on the smallness of the norm of the system-bath interaction. They also prove constructively that given a local noise model, it is possible to simulate an ideal quantum circuit with size L and depth D up to any accuracy, using circuit constructed out of noisy gates from the Boykin set [3] with size $L' = O(L(logL)^a)$ and depth $D' = O(D(logD)^b)$, where a and b are constants that depend on the error correction code that we choose and the design of the fault-tolerant architecture, in addition to more assumptions [1, 2]. These two results combined give us a fault-tolerant threshold theorem for non-Markovian noise, provided that the strength of the effective local noise model is smaller than a positive number that depends on the fault-tolerant architecture we choose. However the ideal measurement process may involve a strong system-bath interaction which necessarily gives a local noise model of large strength. We refine the reduction of the non-Markovian noise model to the local noise model such that this need not be the case, provided that system-bath interactions from the non-ideal operations is sufficiently small. We make all assumptions that [1, 2] has already made, in addition to a few more assumptions to obtain our result. We also give two specific instances where the norm of the fault gets suppressed by some paramater other than the norm of the system-bath interaction. These include the large ratio of the norm of the ideal Hamiltonian to the norm of the perturbation, and frequency of oscillation of the perturbation. We hence suggest finding specific phenomenological models of noise that exhibit these properties.

Acknowledgements

I want to thank my supervisor Debbie Leung for her effective supervision and greatest help throughout all my time in Waterloo. She has helped me learn much about research and this thesis would not have been possible without her guidance. I also like to thank our collaborator Man-Hong Yung, especially for patiently showing me how he formulates problems in physics. I have found that very inspiring.

I also like to thank John Preskill for introducing me to fault-tolerant quantum computing. Special thanks goes to Ian Goulden for teaching me algebraic combinatorics. I have found the philosophy of doing combinatorics taught by Ian to be immensely inspiring and helpful in tackling the mathematical problems at hand. I also want to thank my readers Ben Reichardt and Andrew Childs.

I also want to thank the Mike and Ophelia Lazaridis fellowship for finanical support. I am also very grateful to my friends who have offered me encouragement throughout my studies. Most of all, I want to thank my parents for their unconditional support through all my ups and downs.

Contents

Chapter 1

Introduction

It has been more than two decades since Feynman [4, 5] among others [6, 7] envisioned the concept of the quantum computer. From Feynman's point of view, a quantum computer is a machine that runs using the laws of quantum mechanics and its objective is to simulate quantum mechanics. Such a machine, if built, will help us to better understand a whole range of physical phenomenon such as simulating quantum field theories [8, 9, 10], estimating the eigenvalues and eigenvectors of physical Hamiltonians [11, 12] and simulating many body quantum systems [13, 14]. In addition, algorithms designed to make fundamental use of quantum mechanics have been shown to solve certain problems more efficiently than their classical counterparts, such as factoring [15] and searching for marked items in a large database [16]. The RSA crytosystem [17], on which most transactions in online banking is based on [18], works based on the assumption that factoring composite numbers with large prime factors is hard. Hence a quantum computer can compromise the security of the RSA crytosystem [15], and this is one of the many reasons for the interest in the field of quantum computing.

Nuclear magnetic resonance quantum computers have been built to handle twelve qubits [19], and it has earlier been experimentally demonstrated that Shor's factoring algorithm can factor the composite number 15 into 3 and 5 [20]. However if one intends to tackle more complex problems, we need to be able to manipulate thousands of qubits reliably and without an exponential requirement of classical resources in terms of number of qubits used. Decoherence is a big problem in quantum computing. The field of fault tolerant quantum computation studies the conditions under which it is possible to build large scale quantum computers that perform arbitarily reliably without an excessive use of resources in a non-excessively noisy environment and has

been well studied [21, 1, 22, 23, 24, 25, 26]. Although there are many different models of quantum computation, we restrict our attention to the circuit model, where the quantum computer is built using single or two qubit gates.

Consider the instance in which we have a stochastic noise model for errors occuring in our quantum computer, that is, its components fail independently of each other with a maximum failure probability of $p_{\max}$. Then it has been shown that one can design the architecture of the quantum computer such that if $p_{\max} \leq p_{\text{thres}}$, where p_{thres} can be evaluated, then reliable large scale quantum computation is possible. However from the experimentalist's point of view, the assumption that the noise model is Markovian is unsatisfactory, because it can be experimentally verified that failure of the components of a quantum computer are often correlated with one another in space and in time. Hence there is a need to know the conditions that a non-Markovian noise model must satisfy in order to perform reliable large scale quantum computation.

The authors of [1] have provided an excellent framework from which fault tolerant quantum computation can be studied. They introduced a noise model which they call the **local noise model** with noise strength $\eta \in \mathbb{R}^+$. They prove that, among many other assumptions, provided that the η for the local noise model is less than some positive real number that can be computed, fault tolerant quantum computation is possible. The most important of these assumptions is that η does not scale with the number of qubits used or time required in the quantum computation. Hence if we use the framework of [1], all we need to do is to reduce a physical noise model to a local noise model parametrized by η. For example in the case of the stochastic noise models, $\eta = \sqrt{p_{\max}}$. In the case where the noise model is non-Markovian but local as defined in Chapter 3, then we can also reduce this noise model to a local noise model with a bound for η in Chapter 3.

We observe that when the measurement of qubits in our quantum computer strongly couples the measured qubits to the environment (sharp measurements), the upper bound on η (Lemma 2 of chapter 3) becomes practically impossible to satisfy. We emphasize that sharp measurements do occur in many experimental realizations for quantum computation. A simple example is that of the measurement of a qubit in the ion-trap model of quantum computation [27]. In that scenario, we force a qubit to interact with steady stream of photons emitted from a laser source. The frequency of the laser is chosen such that if the qubit is in its logical 'one' state, it will undergo thousands of Rabi oscillations and will spontaneously emit photons of a known frequency where these emitted photons may be detected and are usually destroyed upon detection. If the qubit is in its 'zero' state, then the interaction of the

steady stream of photons with the qubit is negligible and the qubit will essentially be unperturbed. Hence we wish to take into account the problem of sharp measurement. To resolve this issue, we model the measurement process more carefully in Chapter 6 to obtain a tighter upper bound for the η we obtain by reducing our non-Markovian noise model to the local noise model.

In Chapter 2 we introduce basic mathematical notations, terminology and conventions. In Chapter 3, we study quantum dynamics in closed quantum systems from the Hamiltonian evolution point of view. We use the definition of faults and a fault path expansions from [2, 1]. We then explicitly demonstrate how to upper bound the norm of faults. This will be essential in the reduction of non-Markovian noise models to the local noise models. In this chapter we work out in greater detail claims that have been made in [1] and reorganized much of its material on the non-Markovian noise model.

In Chapter 4 we introduce the nonmenclature used in the circuit model for quantum computation that is relevant to a particular method of constructing fault tolerant quantum circuits. We introduce the notion of concatenation and fault tolerant gadgets from existing literature [21, 1, 2]. In Chapter 5 we define the local noise model which was introduced by [1, 2], and elaborate on its consequences. We reproduce the proof in [1, 2], that a local noise model, together with concatenated quantum circuits built out of fault tolerant gadgets can perform reliably with a modest resource overhead. In Chapter 6 we explicitly describe how non-Markovian noise where sharp measurement is taken into account can be modelled, and we again reduce it to a local noise model.

We believe that it is possible for a non-Markovian noise model to be reduced to a local noise model with strength much smaller than what has been done so far [2, 1], provided that our non-Markovian model is dominated by high-frequency time dependent noise terms, and anticommuting noise terms that have norm small compared to the ideal Hamiltonian. Also we believe that time-dependent anticommuting noise terms of high frequency should have little effect on the evolution. An open problem is to come up with a phenomenological noise model incorporating these features that also makes physical sense, and to demonstrate how this particular phemonological noise model can be reduced to a local noise model with a strength that is less than what would have been obtained if we use the methods of [1, 2].

Although the spin-boson model [28] has been widely used in the physics literature to study non-Markovian noise, it is not easy to reduce it to a useful local noise model [21]. This is another reason for our suggestion of formulating an alternative non-Markovian noise model that incorporates the above features in hope that it will be

analytically more tractable to study. We also study two examples to illustrate that the supremum norm of the system-bath interaction is not the only parameter that size of the fault can depend on, and we hope come up with a more general noise model that illustrates this property.

In this thesis, we provide two examples where the fault is small even when the norm of the system-bath interaction becomes arbitrarily large. In particular, we study two toy models in which the norm of the fault goes to zero as some parameter in the toy model goes to infinity. We demonstrate a special instance of time-independent anticommuting noise to show explicitly how the norm of the fault goes to zero as the ideal Hamiltonian norm becomes very large.

The other toy model is the case where the ideal Hamiltonian is time-independent, and the perturbing Hamiltonian is time-dependent, commutes with the ideal Hamiltonian and oscillates sinusoidally with respect to time with a known frequency ω. Then we can show that provided the norm of the perturbation grows like $o(\omega)$, the norm of the resulting fault goes to zero as ω goes to infinity. This result is an example of having a small fault even when the norm of the perturbation is unbounded.

Chapter 2

Preliminaries

In this thesis, we assume that the postulates of quantum mechanics hold, and not state the axioms of quantum mechanics formally here. We refer the interested reader to [29] for example. Von Neumann showed that the formalism of quantum mechanism can be described using the theory of Hilbert spaces, and we follow this approach for the formalism of quantum mechanics. We will not attempt a thorough introduction to the theory of Hilbert spaces, and will only introduce what we need to make this thesis as self contained as possible. Thus we sometimes prove results that have been claimed to be true in the literature, or sometimes prove proven results as an exercises.

We now introduce notations that we will use in the subsequent chapters. First of all, all physical quantities will have SI units. We do not adopt the convention of setting $\hbar = 1$.

A Hilbert space is a complete inner product space. We remind the reader that a vector space is complete if every Cauchy sequence in it converges. Let $\mathcal{H}$ be the Hilbert space of a quantum system, and let us denote the inner product of $\mathcal{H}$ by $(\cdot, \cdot)_{\mathcal{H}}$. Now let V and W be vector spaces. Let $L(V, W)$ denote the set of all linear operators with domain V and range W. We will use $L(V)$ to abbreviate $L(V, V)$. We usually do not specify the inner product of the Hilbert space of our quantum system explicitly. We will also always work with separable Hilbert spaces, so that they will always admit countable bases.

Now let the set of all elements in $\mathcal{H}$ with norm of 1 be defined as $S(\mathcal{H})$, which we call the set of all pure states of $\mathcal{H}$. The space $\mathcal{H}^* := L(\mathcal{H}, \mathbb{C})$ is called the dual space of $\mathcal{H}$ and is denoted by $\mathcal{H}^*$, and the elements of $\mathcal{H}^*$ are called continuous linear functionals. There are two ways to represent quantum states of a quantum

system with Hilbert space $S(\mathcal{H})$. The first approach is to represent a quantum state of quantum system $\mathcal{H}$ as an element of $\mathcal{H}$. The second approach is to represent a quantum state of a quantum system $\mathcal{H}$ as an element of $L(\mathcal{H})$ which we will call a density operator. We will define the set of density operators in the next paragraph. We will use the Dirac bra-ket notation to represent elements of $\mathcal{H}$ and $\mathcal{H}^*$. For example elements of $\mathcal{H}$ will always be written in the form $|\cdot\rangle$ and elements of $\mathcal{H}^*$ will always be written in the form $\langle\cdot|$. Suppose that we have $\langle f| \in \mathcal{H}^*$ and $|g\rangle \in \mathcal{H}$. Then the operator denoting the linear functional $\langle f|$ operating on $|g\rangle$ will be written as $\langle f|g\rangle$. Now by the Riesz Lemma [30], for each $\langle f| \in \mathcal{H}^*$ there is a unique $|f\rangle \in \mathcal{H}$ such that $\langle f|x\rangle = (|f\rangle, |x\rangle)_{\mathcal{H}}$ for all $|x\rangle \in \mathcal{H}$. Thus $\langle f|x\rangle$ can be interpreted as the inner product on $\mathcal{H}$ of the $|f\rangle \in H$ corresponding to $\langle f| \in \mathcal{H}^*$ with $|x\rangle \in \mathcal{H}$. From the fact that $\langle f| \in L(\mathcal{H}, \mathbb{C})$ it follows that $\langle f|g\rangle \in \mathbb{C}$. The notation $|f\rangle\langle g|$ will be understood to be an operator takes any $|h\rangle \in \mathcal{H}$ to $|f\rangle\langle g|h\rangle$ which is an element of $\mathcal{H}$.

Now consider some $|\psi\rangle \in S(\mathcal{H})$. We define the density operator corresponding to $|\psi\rangle$ to be $\rho_\psi = |\psi\rangle\langle\psi|$, and we define ρ_ψ to be a pure state. Now let us consider an ensemble of pure states $\{\rho_{\psi_i}\}_{i=1}^k$ where each pure state ρ_{ψ_i} is defined as $\rho_{\psi_i} = |\psi_i\rangle\langle\psi_i|, |\psi_i\rangle \in S(\mathcal{H})$ and occurs with probability $p_i > 0$ such that $\sum_{i=1}^k p_i = 1$ for all $i \in \{1, ...k\}$ and for some $k \in \mathbb{Z}^+$. This ensemble of pure states can be represented as a density operator $\rho_{\{p_i,\psi_i\}_{i=1}^k} = \sum_{i=1}^k p_i \rho_{\psi_i}$. We emphasize that as long as a the rank of our density operator is not 1, our density operator is not uniquely associated with an ensemble of states. We define the set of all density operators acting on $\mathcal{H}$ as $D(\mathcal{H})$.

We observe that the special case where the Hilbert space $\mathcal{H}$ is finite dimensional offers a large amount of simplification of the formalism that we have introduced above. When $\mathcal{H}$ is finite dimensional, we will always assume that $\mathcal{H}$ is a complex Euclidean space. In this case, $D(\mathcal{H})$ will be the set of all positive semidefinite operators on $\mathbb{C}^n$ with a trace of 1. However we also want to be able to allow for the case where the case where our quantum system has infinite degrees of freedom, which is often the case for environment-quantum computer interactions.

We will often talk about Hamiltonians defined with respect to a Hilbert space $\mathcal{H}$ that corresponds to a quantum system. Hamiltonians will always be Hermitian operators on $\mathcal{H}$ and have units of energy. We are to interpret the Hamiltonian as an observable of a given quantum system that corresponds to the measured energy of the system. The Hamiltonian is also to be interpreted as the generator of the dynamics of a closed quantum system. This will be explained further in the next chapter. If two Hamiltonians H_1 and H_2 in $L(\mathcal{H})$ are related by the equation $H_1 = \gamma\mathbb{1} + H_2$

where $\mathbb{1}$ is the identity operator on $\mathcal{H}$ and $\gamma \in \mathbb{R}$, then we say that H_1 and H_2 are equivalent Hamiltonians. If it is not otherwise stated, we always assume that the Hamiltonian of a system is time independent. If a Hamiltonian H is time dependent, we will denote it by $H(t)$ where the variable t usually is a variable that representing time. We will sometimes use t to refer to the number of errors a quantum error correction code can correct, and this should be clear from the context.

Now let us define a quantum operator on a separable Hilbert space $\mathcal{H}$. Any quantum operation can be described in the form as given by Kraus' representation theorem [31], which is also known as the operator sum representation of quantum operations. We will prefer to use the notation as given in [32]. Essentially the theorem states the following. Suppose that we have a quantum system with Hilbert space $\mathcal{H}$ and our state is the density operator $\rho \in D(\mathcal{H})$. Suppose that our quantum operation is $\mathcal{E}$ and maps $D(\mathcal{H})$ onto $D(\mathcal{H})$. Then we can always write

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger$$

where $E_k \in L(\mathcal{H})$, $\dagger$ is the adjoint operator, and

$$\sum_k E_k^\dagger E_k = \mathbb{1}$$

where $\mathbb{1}$ is the identity operator on $\mathcal{H}$. The second condition represents the requirement that our quantum operations are trace preserving. The summation is over a countable set since we have assumed we only work with separable Hilbert spaces. We call the set of operators $\{E_k\}$ corresponding to the quantum operation $\mathcal{E}$ the set of Kraus operators corresponding to quantum operation $\mathcal{E}$.

Let us now introduce some notation relevant to quantum error correction codes. Let us define the set $\{\mathtt{X}, \mathtt{Y}, \mathtt{Z}\}$ to be the set of Pauli matrices and let $\mathtt{I} = \mathtt{X}^2 = \mathtt{Y}^2 = \mathtt{Z}^2$ where

$$\mathtt{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \mathtt{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \mathtt{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \mathtt{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We define the commutator of $A, B \in L(\mathcal{H})$ as $[A, B] := AB - BA$.

Chapter 3

Quantum Dynamics

We assume that any realization of a quantum computer has to obey the postulates of quantum mechanics. Thus understanding how quantum systems evolve with respect to time within the context of quantum mechanics is paramount to understanding how quantum computers might work. We define the study of how quantum systems evolve with time as quantum dynamics, with the word quantum meaning that systems we deal with obey the postulates of quantum mechanics, and the word dynamics meaning that we are interested in their time evolution.

There are two different approaches to studying quantum dynamics – the open system and the closed system approach. A closed quantum system is a quantum system that does not interact with its environment and an open quantum system is one that does. We will only discuss the dynamics of closed quantum systems. The dynamics of a closed quantum system depends only on its Hamiltonian. It is also possible to describe quantum dynamics of a system that interacts with an environment without explicitly dealing with the Hilbert space of the environment, and we refer the reader interested in this approach to [33]. We will only state the mathematical properties that the Hamiltonian of a quantum system must satisfy. The derivation of the Hamiltonian of quantum systems given the knowledge of its physical properties is beyond the scope of this thesis, and we refer the reader interested in derivation of Hamiltonians for simple physical models to [29].

3.1 Time Evolution of Closed Quantum Systems

In this section we will describe how quantum states in a closed quantum system evolve with respect to time. The material covered in this section can be found in most introductory quantum mechanics textbooks.

Let the Hilbert space of our closed quantum system be $\mathcal{H}$. Let the Hamiltonian of the quantum system be a Hermitian operator $H \in L(\mathcal{H})$. This means that for all $|f\rangle, |g\rangle \in H$, $(H|f\rangle, |g\rangle)_{\mathcal{H}} = (|f\rangle, H|g\rangle)$. Now assume that at time t_0 our quantum system is in a pure state $|\psi(t_0)\rangle$. It is a postulate of quantum mechanics that any pure state in a closed quantum system evolves according to the Schrödinger equation. This implies that for all $t \in \mathbb{R}, |\psi(t)\rangle \in S(\mathcal{H})$,

$$i\hbar \frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle \tag{3.1.1}$$

where $\hbar$ is the reduced Planck constant.

3.1.1 Time Independent Evolution

In the special case where H is time independent, we can solve the differential equation in (3.1.1) and impose the appropriate initial condition to obtain

$$|\psi(t_0 + t)\rangle = \exp[-iHt/\hbar]|\psi(t_0)\rangle \tag{3.1.2}$$

for all $t \in \mathbb{R}, t \geq 0$. Observe that $\exp[-iHt/\hbar]$ is a unitary operator and is hence an isometry on $S(\mathcal{H})$. It is also independent of the initial state of our quantum system. Using the above equation we can deduce that

$$\langle\psi(t_0 + t)| = \langle\psi(t_0)| \exp[iHt/\hbar] \tag{3.1.3}$$

and hence

$$|\psi(t_0 + t)\rangle\langle\psi(t_0 + t)| = e^{-iHt/\hbar}|\psi(t_0)\rangle\langle\psi(t_0)|e^{iHt/\hbar}. \tag{3.1.4}$$

This implies that if the density operator at time t_0 is $\rho(t_0)$, for all $t \geq 0$ we have

$$\rho(t_0 + t) = e^{-iHt/\hbar}\rho(t_0)e^{iHt/\hbar}. \tag{3.1.5}$$

3.1.2 Time Dependent Evolution

In the case where H is time dependent, we will write H as $H(t)$ in this subsection to emphasize its time dependence for $t \in \mathbb{R}$. Then if our initial pure state at time t_0 is $|\psi(t_0)\rangle \in \mathcal{H}$, then we have

$$|\psi(t_0 + t)\rangle = U|\psi(t_0)\rangle \tag{3.1.6}$$

where

$$U = \lim_{N\to\infty} \prod_{k=1}^{N} \exp[-i\hbar^{-1}H(t_0 + kt/N)t/N],$$

provided $H(t)$ is smooth [29]. We call U the propagator for the time dependent evolution. Now suppose that $H(t)$ can be written as a sum of Hermitian operators, that is, $H(t) = \sum_\sigma H_\sigma(t)$ where $H_\sigma(t) \in L(\mathcal{H})$ for all σ and $t \in \mathbb{R}$. Then by the Trotter product formula [30], we have

$$\exp[-i\hbar^{-1}H(t_0 + kt/N)t/N] = \lim_{M\to\infty} \prod_{j=1}^{M} \prod_{\sigma} e^{-i\hbar^{-1}H_\sigma(t_0+\frac{kt}{N})t/NM} \tag{3.1.7}$$

Thus the propagator U can be rewritten as

$$U = \lim_{N\to\infty} \prod_{k=1}^{N} \lim_{M\to\infty} \prod_{j=1}^{M} \prod_{\sigma} e^{-i\hbar^{-1}H_\sigma(t_0+kt/N)t/NM}.$$

It can be shown that the above equation can be simplified to

$$U = \lim_{N\to\infty} \prod_{k=1}^{N} \prod_{\sigma} e^{-i\hbar^{-1}H_\sigma(t_0+kt/N)t/N}.$$

Observe that in the case where $H(t)$ is time independent, the propagator U simplifies to

$$U = \lim_{N\to\infty} \prod_{k=1}^{N} \exp[-i\hbar^{-1}Ht/N] = \exp[-iHt/\hbar]$$

which is precisely the unitary operator that we have in the previous subsection. The decomposition of our propagator into an infinite product of infinitesimal unitary evolutions has widespread use – for example it is used to derive the equivalence of the Feynman path integral with the propagator derived from the Schrödinger

equation [29]. In particular we will use this decomposition repeatedly throughout this chapter.

The problem that we will later study in great detail has the following structure. We suppose that our Hamiltonian $H(t)$ can be expressed as a sum of two Hamiltonians H_0 and $H_1(t)$, where H_0 is time independent and $H_1(t)$ is time dependent. So we have $H(t) = H_0 + H_1(t)$. The interpretation will be that H_0 is the Hamiltonian that we wish to implement and $H_1(t)$ is the perturbation to our ideal Hamiltonian. We will be interested in evaluating

$$\Delta = \lim_{N\to\infty} \prod_{k=1}^{N} \exp[-i\hbar^{-1}(H_0 + H_1(kt/N))t/N] - \exp[-i\hbar^{-1}H_0 t].$$

3.2 Non-Markovian Noise

Suppose that we have a closed quantum system with Hilbert space $\mathcal{H}$ which is made up of a **system** with Hilbert space $\mathcal{S}$ and a **bath** with Hilbert space $\mathcal{B}$. Suppose that $\mathcal{H} = \mathcal{S} \otimes \mathcal{B}$. Let $\mathbb{1}_B$ denote the identity operator on $\mathcal{B}$ and $\mathbb{1}_S$ denote the identity operator on $\mathcal{S}$. Now any Hamiltonian H in $\mathcal{H}$ can be written in the form

$$H = S \otimes \mathbb{1}_B + \mathbb{1}_S \otimes B + \sum_k S_k \otimes B_k$$

where S and S_k are Hermitian operators in $L(\mathcal{S})$ and S_k is not proportional to $\mathbb{1}_S$ for all k, B and B_k are Hermitian operators in $L(\mathcal{B})$ and B_k is not proportional to $\mathbb{1}_B$ for all k. Now we define $H_S := S \otimes \mathbb{1}_B$, $H_B := \mathbb{1}_S \otimes B$ and $H_{SB} := \sum_k S_k \otimes B_k$. Then H can be written as

$$H = H_S + H_B + H_{SB}.$$

We call H_S the system Hamiltonian, H_B the bath Hamiltonian, and H_{SB} the system bath Hamiltonian. When H_{SB} is zero, we say that the system does not interact with the bath. When H_{SB} is a non-zero operator, we say that the system interacts with the bath. We only demand that $H(t)$ in smooth, and call our noise model the non-Markovian noise model.

As we have seen in the previous section, the dynamics of a closed quantum system is totally described by its propagator. In our current case, the propagator corresponding to an evolution of time $t > 0$ where the initial time is set to $t_0 = 0$ is

$$U = \lim_{N\to\infty} \prod_{k=1}^{N} \exp[-i\hbar^{-1}(H_S(kt/N) + H_B(kt/N) + H_{SB}(kt/N))t/N]$$

Now suppose that we are only interested in the dynamics of the state in $\mathcal{S}$, and that we define the ideal propagator U_{ideal} to correspond to the case where $H_{SB} = 0$ and where the bath Hamiltonian is $H_{B'} \in \mathbb{1}_S \otimes L(\mathcal{B})$. We emphasize that $H_{B'}$ need not be equivalent to H_B. Then $U_{\text{ideal}} = U_S U_{B'}$ where

$$U_S = \lim_{N\to\infty} \prod_{k=1}^{N} \exp[-i\hbar^{-1} H_S(kt/N))t/N]$$

and

$$U_{B'} = \lim_{N\to\infty} \prod_{k=1}^{N} \exp[-i\hbar^{-1} H_{B'}(kt/N))t/N]$$

since $H_{B'}$ and H_S commute by definition. Observe that U_{ideal} depends on our choice of $H_{B'}$. Now observe that $U = U_{\text{ideal}} U_{\text{ideal}}^{-1} U$. Also observe that we can always write $U_{\text{ideal}}^{-1} U = \mathbb{1}_S \otimes V + F$ where $V \in L(\mathcal{B})$ and $F \in L(\mathcal{H})$ and $F \notin \mathbb{1}_S \otimes L(\mathcal{B})$. We further demand that V is a unitary operation on $\mathcal{B}$. Thus it follows that we can always write the propagator U as

$$U = U_{\text{ideal}}(\mathbb{1}_S \otimes V + F).$$

We call F a fault. F and V will depend on our choice of $H_{B'}$. If we wish to interpret $U_{\text{ideal}}(\mathbb{1}_S \otimes V)$ as an ideal operation, then we will have to demand that V is a unitary operator on $L(\mathcal{B})$. In this case, V will be the propagator associated with the dynamics of evolution produced by $H_{B'}$. We will often be interested in finding an upper bound on the supremum norm of F. Now define the operator $G \in L(\mathcal{H})$ where $G := U - U_{\text{ideal}}$. In particular, G corresponds to the choice where V is the identity operation on the bath. Since $U = U_{\text{ideal}} \mathbb{1}_S \otimes V + U_{\text{ideal}} F$, it follows that if $V = \mathbb{1}_B$ and $H_{B'} = H_B$, we will have $U_{\text{ideal}} F = G$ which implies that $\|F\| = \|G\|$. Thus in the general case where we do not choose V and $H_{B'}$ apriori, we will have the inequality $\|F\| \leq \|G\|$. Thus in order to find an upper bound for $\|F\|$, it suffices to find an upper bound for $\|G\|$. However, it is not clear if this upper bound is tight.

We can evaluate $\|F\|$ exactly in the general case, although we will not use this form for any of our results. We only include the following calculations for completeness. By rearranging the terms of the above equation, we get

$$U_S U_{B'} F = U - U_S U_{B'}(\mathbb{1}_S \otimes V).$$

Since U_S and $U_{B'}$ are unitary, this implies that

$$\|F\| = \|U - U_S U_{B'}(\mathbb{1}_S \otimes V)\|.$$

We will now proceed to show that we can have an upper bound for $\|F\|$ which is given by the inequality $\|F\| \leq \|G\| \leq \hbar^{-1} \max_{t' \in [0,t]} \|H_{SB}(t')\| t$. A lemma similar to the one that we will present below has been proved in [21]. This lemma essentially says that if our perturbation has a supremum norm at most α and our time evolution is allowed to occur for time t, then the supremum norm of the difference between the actual propagator and the ideal propagator is at most αt. Intuitively, α quantifies the strength of the system-bath interaction. Hence this lemma says that if the product of α and the time t is small, then the system and bath do not interact much and so the system evolution does not deviate too much much from the ideal evolution.

Lemma 1 *Suppose that we have a Hilbert space $\mathcal{H}$ and $U_k, A_k \in L(\mathcal{H})$ where A_k is Hermitian for all $k \in \{1, ..., N\}$. Let t be a positive constant. Let $||A_k|| = \tilde{\alpha}(kt/N) \leq \alpha$ for all $k \in \{1, ..., N\}$, and $\tilde{\alpha}(\cdot)$ be a real nonnegative bounded function with a domain of $[0, t]$ and $\alpha = \sup_{t' \in [0,t]} \tilde{\alpha}(t')$. Define $U := U_N ... U_1$ and $U_{\text{bad}} := U_N e^{-iA_N t/N} ... U_1 e^{-iA_1 t/N}$. Then (i)*

$$\|U - U_{\text{bad}}\| \leq \alpha t. \tag{3.2.1}$$

If $\tilde{\alpha}(\cdot)$ is Riemann integrable, then (ii)

$$\lim_{N \to \infty} ||U - U_{\text{bad}}|| \leq \int_{t'=0}^{t} \tilde{\alpha}(t') dt'. \tag{3.2.2}$$

Proof: Let $V_k = U_k e^{-iA_k t/N}$ for all $k \in \{1, ..., N\}$. Then

$$\begin{aligned} U_{\text{bad}} - U =& V_N V_{N-1} ... V_2 V_1 - U_N U_{N-1} U_{N-2} ... U_2 U_1 \\ =& V_N V_{N-1} ... V_2 V_1 - U_N V_{N-1} V_{N-2} ... V_2 V_1 \\ &+ U_N V_{N-1} ... V_2 V_1 - U_N U_{N-1} V_{N-2} ... V_2 V_1 \\ &\vdots \\ &+ U_N U_{N-1} ... U_2 U_1 - U_N U_{N-1} ... U_2 U_1. \end{aligned}$$

In the second equality, we have rewritten U_{bad} as a telescopic sum. Since V_k is unitary,

$\|V_k\| = 1$ for all $k \in \{1, ..., N\}$ and hence

$$\begin{aligned}\|U_{\text{bad}} - U\| &\leq \sum_{k=1}^{N} \|V_k - U_k\| \\ &= \sum_{k=1}^{N} \|U_k(e^{-iA_k t/N} - 1)\| \\ &\leq \sum_{k=1}^{N} \|A_k t/N\|\end{aligned}$$

If we use the fact that $\|A_k\| \leq \alpha$ for all $k \in \{1, ..., N\}$, then we get

$$\|U_{\text{bad}} - U\| \leq \alpha t$$

which proves the first part of the lemma. If we apply the definition of $\|A_k\|$ and the definition of Riemann integrability, we will obtain the second part of the lemma. ■

Now let $\alpha = \hbar^{-1} \sup_{t' \in [0,t]} \|H_{SB}(t')\|$. Then it follows by direct application of the first result of Lemma 1 that $\|F\| \leq \|G\| \leq \alpha t$. Now if we suppose that $\|H_{SB}(t')\|$ is Riemann integrable on $t' \in [0, t]$ then by the second result of Lemma 1 we get $\|F\| \leq \hbar^{-1} \int_0^t \|H_{SB}(t')\| dt'$. Clearly the latter bound is tighter than the first one because we have made an additional assumption.

We emphasize that the bound of Lemma 1 is by no means tight, even when we assume that $\|H_{SB}(t')\|$ is Riemann integrable on $t' \in [0, t]$. By definition $W := \mathbb{1}_S \otimes V + F$ is a unitary operator, and thus $\|W\| = 1$. In the special case where $V = \mathbb{1}_B$ and $H_{B'} = H_B$, we can arrive at the equation $\|W - \mathbb{1}_S \otimes \mathbb{1}_B\| = \|F\|$. By the triangle inequality for norms, we can get $\|F\| \leq 2$. However observe that if $\hbar^{-1} \int_0^t dt' \|H_{SB}(t')\| > 2$ when either $\min_{t' \in [0,t]} \|H_{SB}(t')\| > 0$ and t becomes very large, or if $\hbar^{-1} \min_{t' \in [0,t]} \|H_{SB}(t')\|$ is infinite, then our inequality $\|F\| \leq \hbar^{-1} \int_0^t dt' \|H_{SB}(t')\|$ will not be useful.

Unfortunately there exist models of non-Markovian noise model where H_{SB} is unbounded. An example is the spin-boson noise model, where the system is a single spin 1/2 particle which couples linearly to a bath of bosonic quantum harmonic oscillators. This phenomenological model is well studied in the physics literature [28], and may be used as a toy model to study non-Markovian noise. The interaction term is unbounded because position and momentum operators are unbounded. In particular, [21] has studied the spin-boson model in the context of fault-tolerant quantum computation.

There are examples where $\|F\|$ is small even when H_{SB} becomes arbitrarily large, and of the goals of this thesis is to understand when this happens. Let the ideal Hamiltonian have a supremum norm of h. We will study two specific examples – anticommuting perturbation with a supremum norm of a, and high frequency commuting perturbation with a supremum norm of c. We show that in a particular instance of anticommuting noise, a time-independent upper bound on $\|F\| \leq O(a/h)$ which is time independent can be obtained. In a particular instance when the perturbation is of the commuting type and of high frequency ω, we can show that $\|F\|$ remains small if $c = o(\omega)$ as ω becomes arbitrary large. These two examples illustrate that the supremum norm of H_{SB} is not the only parameter that $\|F\|$ can depend on, and we hope come up with a more general noise model that illustrates this property.

3.2.1 Anticommuting Noise

We like to bring to the reader's attention a special case where we are able to calculate the effective fault induced by the noise exactly. Although this particular example may seem to be very artificial, we hope that it may give us insight on how to upper bound $\|F\|$ more tightly for the case where H_{SB} has a more general form. The following theorem applies to the scenario in which we have anticommuting noise

Theorem 1 *Suppose that we have a separable Hilbert space $\mathcal{H}$ and have Hermitian operator $H \in L(\mathcal{H})$ and a linear operator $A \in L(\mathcal{H})$ such that $HA = -AH$. Let $H = \sum_m h_m |m\rangle\langle m|$ be the spectral decomposition of H. Let $\mathbb{1}$ be the identity operator on $\mathcal{H}$. Then for $t > 0$,*

$$e^{-i(H+A)t} = \sum_m |m\rangle\langle m| \left(\cos(t\sqrt{h_m^2 \mathbb{1} + A^2}) - i(h_m \mathbb{1} + A)t \,\mathrm{sinc}(t\sqrt{h_m^2 \mathbb{1} + A^2}) \right) \tag{3.2.3}$$

where $\mathrm{sinc}(x) := \sin(x)/x$ *is a formal power series in x.*

We prove this theorem in the appendix.

In the special case where $A^2 = a^2\mathbb{1}, H^2 = h^2\mathbb{1}$ where $a, h > 0$ we can simplify $e^{-i(H+A)t}$ to get

$$\begin{aligned} e^{-i(H+A)t} &= \sum_m |m\rangle\langle m| \left(\mathbb{1}\cos(t\sqrt{h^2+a^2}) - i(h_m\mathbb{1} + A)t\,\mathrm{sinc}(t\sqrt{h^2+a^2}) \right) \\ &= \mathbb{1}\cos(t\sqrt{h^2+a^2}) - it(H+A)\,\mathrm{sinc}(t\sqrt{h^2+a^2}) \\ &= \mathbb{1}\cos(h\tau) - i\frac{H+A}{\sqrt{h^2+a^2}}\sin(h\tau). \end{aligned} \tag{3.2.4}$$

where $\tau = \sqrt{1+(a/h)^2}$. Here we can observe that anticommuting noise shifts the timescale at which the dynamics of H occurs. Now observe that

$$\begin{aligned} \left\| e^{-i(H+A)t} - e^{-iH\tau} \right\| &= \left\| i\sin(h\tau)\left(\frac{H}{h} - \frac{H+A}{\sqrt{h^2+a^2}} \right) \right\| \\ &\leq |\sin(h\tau)| \left(\|H\| \left| \frac{1}{h} - \frac{1}{\sqrt{h^2+a^2}} \right| + \|A\| \frac{1}{\sqrt{h^2+a^2}} \right) \\ &\leq \left| 1 - \frac{1}{1+a^2/h^2} \right| + a/h \end{aligned} \tag{3.2.5}$$

What this implies for us is that in the case where we know what the anticommuting noise is, and when a/h is very small, we can expect the effective fault size to be very small. However if we do not know what the anticommuting noise is, we will not know how the timescale of the dynamics of H will have shifted, and thus the fault size in this case can be large.

We believe that this type of result may also hold for the case where A^2 and H^2 are not necessarily proportional to $\mathbb{1}$, and A is a time-dependent operator oscillating with a very high frequency. This is because we believe that if we expand out F in the form of $D.7$, the coefficients for the higher order spherical Bessel functions will become negligible. It remains an open problem to show that this is indeed the case.

3.2.2 High Frequency Noise

Consider two Hamiltonians $H, P \in L(\mathcal{H})$ where $\mathcal{H}$ is some Hilbert space. Assume that H and P commute and are time independent. Let the Hamiltonian describing the dynamics of our perturbed system be $H_{\text{bad}} = H + P\cos(\omega t)$ where $\omega > 0, t \in [0, \tau], \tau > 0$. Let the propagator corresponding to the dynamics generated by H_{bad} from time $t = 0$ to time $t = \tau$ be $U_{\text{bad},\tau}$. Let the ideal Hamiltonian be H and let the

propagator corresponding to the dynamics generated by H from time $t = 0$ to time $t = \tau$ be $U_{\text{ideal},\tau}$. We will now evaluate $F_p := \|U_{\text{bad},\tau} - U_{\text{ideal},\tau}\|$.

Observe that $U_{\text{ideal},\tau} = e^{-iH\tau/\hbar}$. We will now solve for $U_{\text{bad},\tau}$ by first using the Trotter decomposition followed by direct use of the Schrödinger equation. Since the perturbation $P\cos(\omega t)$ commutes with H for all t, $U_{\text{bad},\tau} = U_{\text{ideal},\tau}U_{P,\tau}$ where $U_{P\cos(\omega t),\tau}$ is the propagator corresponding to the dynamics generated by the Hamiltonian $P\cos(\omega t)$. Thus $\|U_{\text{bad},\tau} - U_{\text{ideal},\tau}\| = \|U_{P\cos(\omega t),\tau} - \mathbb{1}\|$ where $\mathbb{1}$ is the identity operator on $\mathcal{H}$. Now by using the standard procedure of solving separable partial differential equations, and matching to the initial condition at time $t = 0$, we readily obtain

$$U_{P\cos(\omega t),\tau} = e^{-iP\sin(\omega t)t/\hbar}.$$

Observe now that for all $t = 2n\pi/\omega$ for $n \in \mathbb{Z}$, $U_{P\cos(\omega t),\tau} = \mathbb{1}$. Thus we can obtain

$$U_{P\cos(\omega t),\tau} = e^{-iP\sin(\omega t')t'/\hbar}$$

where $t' = \min_{n\in\mathbb{N}}\{t - 2n\pi/\omega\}$. Clearly $0 \leq t' \leq 2\pi/\omega$. Then we can use the first part of Lemma 1 to show that $F_p \leq \|P\|t' \leq 2\pi\|P\|/\hbar\omega$. Hence if $\|P\| = o(\omega)$ as $\omega \to \infty$, $\|F_p\|$ will approach zero as ω becomes very large, even though $\|P\|$ may be very large.

3.3 Fault-Path Expansion [1, 2]

3.3.1 Multiple systems, one time step

Now assume that there are n distinct systems that we wish to manipulate simultaneously from time $t = 0$ to time $t = t_0$. We are studying the dynamics of only one time step as opposed to multiple time steps only for the sake of pedagogy. The arguments that follow can be trivially extended for the case in which we deal with multiple time steps. We label the Hilbert space of each of these systems as $\mathcal{S}_k$ where $k \in \mathcal{J} := \{1, ..., n\}$. We are interested in the dynamics of the system with Hilbert space $\mathcal{S} := \otimes_{k=1}^{n}\mathcal{S}_k$. Let there be a bath with Hilbert space $\mathcal{B}$ such that our system and bath form a closed quantum system. We define $\mathbb{1}_{S_k}$ to be the identity operator on $\mathcal{S}_k$ for all $k \in \mathcal{J}$ and we define $\mathbb{1}_S$ and $\mathbb{1}_B$ to be the identity operators on $\mathcal{S}$ and $\mathcal{B}$ respectively. We define the entire quantum system to be closed with a Hilbert space of $\mathcal{H} = \mathcal{S} \otimes \mathcal{B}$. Assume that the Hamiltonian of the bath is B where B acts trivially on $\mathcal{S}$.

Define $\mathcal{P}(\mathcal{J})$ to be the powerset of $\mathcal{J}$. For some $\sigma \in \mathcal{P}(\mathcal{J})$, let W_σ denote the set of linear operators in $L(\mathcal{H})$ that act trivially on $\mathcal{S}_j$ for all $j \in \sigma$ and non-trivially on $\mathcal{S}_j$ otherwise. Let $H_{SB,\sigma}$ be a Hermitian operator that belongs to the set W_σ for some $\sigma \in \mathcal{P}(\mathcal{J})$. Then we can always decompose any Hermitian operator $H_{SB} \in L(\mathcal{H})$ into the form

$$H_{SB} = \sum_{\sigma \in \mathcal{P}'} H_{SB,\sigma}$$

where $\mathcal{P}'$ is some subset of $\mathcal{P}(\mathcal{J})$.

We define the ideal Hamiltonian for our system to be $H_{\text{ideal}} := B + \sum_{k=1}^{n} H_{S_k}$ where for all $k \in \mathcal{J}$ we have Hermitian $H_{S_k} \in W_{\{k\}}$ acting trivially on the bath $\mathcal{B}$, and $B \in L(\mathcal{H})$ acts trivially on $\mathcal{S}$. Let U_{ideal} be the ideal propagator corresponding to the dynamics due to the ideal Hamiltonian H_{ideal}. Let the ideal propagator corresponding to system k be

$$U_k := \lim_{N \to \infty} \exp[-i\hbar^{-1} H_{S_k}(jt_0/N)t_0/N].$$

Since all of the Hermitian H_{S_k} operators commute with one another for all $k \in \mathcal{J}$, it follows that $U_{\text{ideal}} = \prod_{k=1}^{n} U_k$ where we are free to permute the indices labelling our Hilbert spaces $\mathcal{S}_k$.

Now let the total Hamiltonian describing the dynamics of our entire system be H_{actual} where

$$H_{\text{actual}} = H_{\text{ideal}} + H_{SB}.$$

Now we are in a position to define a **fault path expansion**.

Definition 1 (Fault Path Expansion, one time step) *Assume that we use the notation defined in this current subsection. Let the propagator corresponding to the Hamiltonian H_{actual} be U_{actual} and the propagator corresponding to H_{S_k} be U_k for all $k \in \mathcal{J}$. The* **fault path expansion** *for U_{actual} is the rewriting of U_{actual} in the form*

$$U_{\text{actual}} = \prod_{k=1}^{n} U_k \Delta_k$$

where Δ_k is a unitary operator on $\mathcal{H}$ for all $k \in \mathcal{J}$, and can be written in the form

$$\Delta_k = (V_k + F_k)$$

where $V_k \in W_{\{k\}}, F_k \in L(\mathcal{H})$ for all $k \in \mathcal{J}$. We also require that F_k acts non-trivially on $\mathcal{S}_k$. The word 'expansion' is used because U_{actual}, when written in this form, can be expanded out as a sum of 2^n monomials, each of which we call a **fault path**. *The* **number of faults** *in each fault path is the number of occurences of F_k in the given monomial.*

Clearly a fault path expansion for U_{actual} always exists, as we can trivially choose to have $\Delta_k = U_k^{-1}$ for all $k \in \{1, ..., n-1\}$ and $\Delta_n = U_{\text{actual}}$. However this example is only pedagogical and rather useless; there are other fault path expansions. A useful fault path expansion for U_{actual} should have $\max_{k \in \mathcal{J}} \|F_k\|$ minimized. We can write this problem down as an optimization problem, but we have not pursued this line of thought further.

Fortunately, the authors of [1] do suggest a particular fault-path expansion that we can use to obtain an upper bound on $\max_{k \in \mathcal{J}} \|F_k\|$. The fault-path expansion that they consider is one where they express U_{actual} as

$$U_{\text{actual}} = \lim_{N \to \infty} \prod_{j=1}^{N} \left(\exp[-i\hbar^{-1} H_{\text{ideal}}(jt_0/N)t_0/N] \prod_{\sigma \in \mathcal{P}(\mathcal{J})} (\mathbb{1} - i\hbar^{-1} H_\sigma(jt_0/N)t_0/N)) \right).$$

If a particular fault path in the fault path expansion has a fault in system k, it necessarily needs to have at least one term $-i\hbar^{-1} H_\sigma(jt_0/N)t_0/N)$ inside when $k \in \sigma$. The above equation is a fault path expansion that [1] describes as expanding the fault path in fine grain time steps. Hence the norm of the fault in system k is upper bounded by the sum of the norm of all the fine grain fault paths that contribute to F_k. To evaluate this sum, we can directly use Lemma 1. Hence we obtain the following result that [1] has also shown:

Lemma 2 *Assume that we use all the notation in this subsection. Then for all* $k \in \mathcal{J}$,

$$\max_{k \in \mathcal{J}} \|F_k\| \leq t_0 \hbar^{-1} \left\| \sum_{\sigma \in \mathcal{P}', k \in \sigma} H_{SB,\sigma} \right\|$$

Proof: We have described the proof in the above paragraph.

The above result can be straightforwardly applied to a scenario in which we want to control n systems in d timesteps. This is analogous to executing a quantum algorithm with a depth of d. In this case, we index the jth system at timestep k with the integer $nk + j$, where $j \in \{1, ..., n\}, k \in \{0, ..., d-1\}$ and we redefine $\mathcal{J} := \{1, ..., nd\}$. Then lemma 2 will also hold as shown in [1].

Lemma 2 shows that the non-Markovian noise model can be reduced to a local noise model of strength at most $\eta := t_0 \hbar^{-1} \left\| \sum_{\sigma \in \mathcal{P}', k \in \sigma} H_{SB,\sigma} \right\|$, where a local noise

model introduced in [1, 2] will be defined in Chapter 4. We say that the non-Markovian noise is **local** if it satisfies the additional property where η does not increase as we allow the number of subsystems in $\mathcal{S}$ to increase. This means that faults that act collectively on multiple locations are highly suppressed.

Chapter 4

Fault Tolerant Quantum Computation

We will restrict our attention to only one model of quantum computation – the circuit model of quantum computation. We will properly define what a quantum circuit is after defining some standard terminology used in the quantum computing literature. But first of all, we devote some time to study what qubits are, since most quantum algorithms make extensive use of them. This chaper is a review of the nonmenclature of fault tolerant quantum computing, and we do not contribute any new ideas here.

4.1 Nonmenclature of Quantum Circuits

The fundamental idea behind quantum circuits is to build a quantum computer by assembling basic building blocks that belong to a finite set. Following [2], we call each elementary building block a **location**. In short there are three types of locations – preparations, unitaries and measurements.

4.1.1 Preparations

A qubit is a two-level system with a Hilbert space of $\mathbb{C}^2$ with Euclidean norm. We represent the two distinct states of a qubit in the computation basis formally as $|0\rangle = (1,0)$ and $|1\rangle = (0,1)$ where $|0\rangle, |1\rangle \in \mathbb{C}^2$. We define a **preparation** to be the initialization of a qubit in the state $\alpha|0\rangle + \beta|1\rangle$ for all $\alpha, \beta \in \mathbb{C}$ such that $|\alpha|^2 + |\beta|^2 = 1$. We represent such a operation as a '0-prep'.

4.1.2 Single Qubit Gates

Let the Hilbert space of some qubit be $\mathcal{S} = \mathbb{C}^2$. A **single qubit gate** acting on $\mathcal{S}$ is defined by a unitary operator mapping $\mathcal{S}$ to $\mathcal{S}$. Each single qubit gate is defined by its action on the computation basis states $|0\rangle, |1\rangle$ of the qubits concerned. Now we define a few single qubit gates.

We will call I the identity gate on a qubit. In particular, $I|0\rangle = |0\rangle$ and $I|1\rangle = |1\rangle$. We will call H the Hadamard gate on a qubit. Here we have, $H|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $H|1\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. We will call T the $\pi/8$ gate on a qubit. In particular, $T|0\rangle = |0\rangle$ and $T|1\rangle = e^{i\pi/4}|1\rangle$.

The set of single qubit gates that we will use is $\{I, H, T\}$.

4.1.3 CNOT

A CNOT is a unitary gate acting on two qubits. One of the qubits will be chosen as the control qubit and the other as the target qubit. Let the Hilbert space of the control qubit be $\mathcal{C} = \text{span}\{|0\rangle_c, |1\rangle_c\}$ and the Hilbert space of the target qubit be $\mathcal{T} = \text{span}\{|0\rangle_t, |1\rangle_t\}$. Here we use the lower case roman alphabets c and t to label the basis states of our control and target qubits respectively. Then a CNOT with a control on $\mathcal{C}$ and with target $\mathcal{T}$ is a unitary operation mapping $\mathcal{C} \otimes \mathcal{T}$ to $\mathcal{C} \otimes \mathcal{T}$ such that the following equalities are satisfied.

$$\begin{aligned}
\text{CNOT}|0\rangle_c \otimes |0\rangle_t &= |0\rangle_c \otimes |0\rangle_t \\
\text{CNOT}|0\rangle_c \otimes |1\rangle_t &= |0\rangle_c \otimes |1\rangle_t \\
\text{CNOT}|1\rangle_c \otimes |0\rangle_t &= |1\rangle_c \otimes |1\rangle_t \\
\text{CNOT}|1\rangle_c \otimes |1\rangle_t &= |1\rangle_c \otimes |0\rangle_t
\end{aligned}$$

4.1.4 Single Qubit Measurement

Before we can define what a measurement on a single qubit is, we will review the definition of a density matrix corresponding to a qubit. A qubit that is a pure state with a Hilbert space of $\text{span}\{|0\rangle, |1\rangle\}$ and can be represented by $a|0\rangle + b|1\rangle$ where $a, b \in \mathbb{C}$ such that $|a|^2 + |b|^2 = 1$. The corresponding density matrix for this pure state is $(a|0\rangle + b|1\rangle)(a\langle 0| + b\langle 1|)$. In general, a qubit can be in a probabilistic ensemble of k pure states, each pure state occuring with probability p_k where $k \geq 1$. If we

denote the density matrix of each pure state as ρ_i where $i \in \{1, ..., k\}$, then we define the density matrix of our ensemble of pure states to be $\rho := \sum_{i=1}^{k} p_i \rho_i$.

We define the measurement of a qubit with a density matrix ρ in the computation basis to be an operator that projects the state of our qubit to either $|0\rangle$ or $|1\rangle$, with probabilities trace$(\rho|0\rangle\langle 0|)$ and trace$(\rho|1\rangle\langle 1|)$ respectively, where 'trace' is the trace operator that acts on the space of linear operators on the Hilbert space of our qubit. We will denote a single qubit measurement by a '0-meas'.

4.1.5 Locations

We define a **location** to be any element of the set $\mathcal{L} = \{$CNOT, 0-meas, 0-prep, $I, H, T\}$. Observe that $\{H, T,$ CNOT$\}$ is the Boykin set [3]. We will assume that the implementation of each location takes a time of exactly t_0 seconds where $0 < t_0 < \infty$. We define the amount of time needed to implement each location as a **timestep**.

4.2 Quantum Algorithm

We will restrict our attention to quantum algorithm implemented using quantum circuits, and hence we use the terms 'quantum circuit' and 'quantum algorithm' interchangably. We further restrict ourselves to the case where our quantum circuit must be constructed out of elements of the set $\mathcal{L}$.

Now we will define a quantum algorithm more formally. Let us first denote a quantum algorithm by $\mathcal{Q}$. We say that $\mathcal{Q}$ is ideal if it is built out of ideal locations. By convention, we assume that our ideal quantum algorithm $\mathcal{Q}$ always has the following form. First we prepare n-qubits, each in the $|0\rangle$ state. We define the set of these n-qubits to be our **system** with Hilbert space $\mathcal{S}$. We are to implement a unitary operator $U \in L(\mathcal{S})$ on our system $\mathcal{S}$. U will be implemented as a finite product of unitaries having the form $U = \prod_{\ell=1}^{L} U_\ell$ and $L \geq 2n$. ℓ is really an index that labels each component in the quantum circuit for $\mathcal{Q}$. For all $\ell \in \{n+1, ..., L-n\}, U_\ell$ either acts non-trivially on exactly one or two qubits in $\mathcal{S}$ and trivially on everything else or acts trivially on $\mathcal{S}$, and in this case ℓ labels our unitary gates. For $\ell \in \{1, ..., n\} \cup \{L-n+1, ..., L\}$, $U_\ell = I_\mathcal{S}$ where $I_\mathcal{S}$ is the trivial operation on $\mathcal{S}$. Here ℓ labels our preparations for $|0\rangle$ and measurements in the computation basis. We call the quantum circuit element associated with ℓ **location** ℓ. Assume that each **location** takes one timestep to execute. We say that the depth D of $\mathcal{Q}$ is the

minimum number of timesteps needed to implement U. By convention, we have $L = Dn$, and this reflects the fact that we count identity gates on our qubits as gates that we have to implement. After U is implemented, we measure each qubit in $\mathcal{S}$ in the computation basis $\{|0\rangle, |1\rangle\}$. Our quantum algorithm outputs measurement outcomes on on a chosen subset $\mathcal{M}$ of our system qubits that follows the probability distribution $q = \{q_i\}_{i=1}^{2^{|\mathcal{M}|}}$.

4.3 Encoded Qubits and Gates

We will first define a $((n, M))$ quantum code for $n, M \in \mathbb{Z}^+$. Suppose that we have n qubits with Hilbert space $\mathcal{H}$. We define a set of n-qubit states $\mathtt{C} := \{|\overline{0}\rangle, ..., |\overline{M-1}\rangle\}$ to be a $((n, M))$ code if $|\overline{k}\rangle \in S(\mathcal{H})$ and $\langle\overline{k}|\overline{j}\rangle = \delta_{jk}$ for all $j, k \in \{0, ..., M-1\}$ where δ_{jk} is the Kronecker delta function. We wish to define the notion of a codespace with respect to an element of $\mathcal{H}$ and an element of $D(\mathcal{H})$ separately. We say that $|\psi\rangle \in \mathcal{H}$ is in the codespace of $\mathtt{C}$ if $|\psi\rangle \in \mathcal{C}_\mathtt{C} := \{x \in \mathcal{H} : x = \sum_{k=0}^{M-1} \alpha_k |\overline{k}\rangle, \sum_{k=0}^{M-1} |\alpha_k|^2 = 1, \alpha_k \in \mathbb{C} \text{ for all } k \in \{0, ..., M-1\}\}$. We say that $\rho \in D(\mathcal{H})$ is in the codespace of $\mathtt{C}$ if $\rho \in D(\mathcal{C}_\mathtt{C}) := \{x = \sum_k p_k |\psi_k\rangle\langle\psi_k| : |\psi_k\rangle \in \mathcal{C}_\mathtt{C}, p_k \geq 0 \ \ \forall k, \sum_k p_k = 1\}$.

Now we define the projector for a quantum code $\mathtt{C}$ to be $P_\mathtt{C} := \sum_{k=0}^{M-1} |\overline{k}\rangle\langle\overline{k}|$. We say that $\mathtt{C}$ can correct for errors induced by a set Kraus operators $\{E_i\} \subset L(\mathcal{H})$ if the quantum error correction criterion [34, 35] is satisfied, that is

$$P_\mathtt{C} E_i E_j P_\mathtt{C} = \alpha_{ij} P_\mathtt{C}$$

for all $E_i, E_j \in \{E_i\}$ and the matrix corresponding to α_{ij} is a Hermitian matrix. By linearity of quantum operations, if $\{E_i\}$ is a correctible set of errors, then $\{\sum_j \beta_{ij} E_j\}$ is also a correctible set of errors. Let us define the weight of a unitary operator $U \in L(\mathcal{H})$ to be the number of qubits it acts non-trivially on and denote it by $wt(u)$. Let $S_{jk} \in L(\mathcal{H})$ be the set of all unitaries that take $|\overline{j}\rangle$ to $|\overline{k}\rangle$. It is almost always the case that S_{jk} is not a singleton. From the quantum error correction criterion, it is natural to define the distance of the code $\mathtt{C}$ to be the d where

$$d := \min\{wt(U) : U \in S_{jk}, j \in \{0, ..., M-1\}, k \in \{0, ..., M-1\}, j \neq k\}.$$

Now let us consider some density operator on n qubits $\overline{\rho} \in D(\mathcal{H})$. Let us define $d_{\overline{\rho}}$ to be the distance of $\overline{\rho}$ from the codespace $D(\mathcal{C}_\mathtt{C})$. $d_{\overline{\rho}}$ is the minimum $d'_{\overline{\rho}}$ for which the following equality can be satisfied:

$$\rho = \sum_k p_k U_k |\overline{k}\rangle\langle\overline{k}| U_k^\dagger$$

where $|\overline{k}\rangle \in \mathcal{C}_{\mathtt{C}}$, $p_k \geq 0$, $\sum_k p_k = 1$, $U_k \in L(\mathcal{H})$ is unitary, $wt(U_k) \leq d'_{\overline{\rho}}$ for all $k \in \{0, ..., M-1\}$. We say that a quantum error correction code can correct $t = \lfloor (d-1)/2 \rfloor$ errors.

Now we are finally in the position to define encoded locations, which we call 1-Ga, which is the short form for 1-gadget.

Since we are only interested in encoding 1 qubit, we will assume that $\mathtt{C}$ is a $((n, 2))$ quantum code. The encoded operation corresponding to the preparation of a qubit in the $|0\rangle$ state will be called a 1-prep, which we define to initialize a n-qubit state in the $|\overline{0}\rangle$ state. Let $\overline{\rho} \in D(\mathcal{C}_{\mathtt{C}})$ and $\mathcal{E}$ be a quantum operation. A 1-EC is defined to map any state $\mathcal{E}(\overline{\rho})$ that has a distance no more than t to the state $\overline{\rho}$. We define an ideal-decoder to map any state $\mathcal{E}(\overline{\rho})$ that has a distance no more than t to the state ρ where ρ is the density matrix for a single qubit corresponding to $\overline{\rho}$ which is the encoded version. In other words, $\langle k|\rho|j\rangle = \langle \overline{k}|\overline{\rho}|\overline{j}\rangle$ for all $j, k \in \{0, 1\}$. The encoded measurement is a 1-meas, which is the encoded version of the single qubit measurement.

The encoded Hadamard gate denoted by $\overline{H}$ is defined to be a linear operator on the Hilbert space of n-qubits, satisfying the relations $\overline{H}|\overline{0}\rangle = (|\overline{0}\rangle + |\overline{1}\rangle)/\sqrt{2}$ and $\overline{H}|\overline{1}\rangle = (|\overline{0}\rangle - |\overline{1}\rangle)/\sqrt{2}$. The encoded pi/8 gate denoted by $\overline{T}$ is defined to be a linear operator on the Hilbert space of n-qubits satisfying the relations $\overline{T}|\overline{0}\rangle = |\overline{0}\rangle$ and $\overline{T}|\overline{1}\rangle = e^{i\pi/4}|\overline{1}\rangle$. The encoded identity gate denoted by $\overline{I}$ is defined to be a linear operator on the Hilbert space of n-qubits satisfying the relations $\overline{I}|\overline{0}\rangle = |\overline{0}\rangle$ and $\overline{I}|\overline{1}\rangle = |\overline{1}\rangle$. The encoded CNOT denoted by $\overline{\text{CNOT}}$ is defined to be a linear operator on the Hilbert space of $2n$-qubits. Let the Hilbert space of the first n qubits be $\mathcal{C} = \text{span}\{|\overline{0}\rangle_c, |\overline{1}\rangle_c\}$ and the Hilbert space of the next n qubits be $\mathcal{T} = \text{span}\{|\overline{0}\rangle_t, |\overline{1}\rangle_t\}$. Here we use the lower case roman alphabets c and t to label the basis states of our control and target qubits respectively. Then a $\overline{\text{CNOT}}$ with a control on $\mathcal{C}$ and with target $\mathcal{T}$ is a unitary operation mapping $\mathcal{C} \otimes \mathcal{T}$ to $\mathcal{C} \otimes \mathcal{T}$ such that the following equalities are satisfied.

$$\begin{aligned}
\overline{\text{CNOT}}|\overline{0}\rangle_c \otimes |\overline{0}\rangle_t &= |\overline{0}\rangle_c \otimes |\overline{0}\rangle_t \\
\overline{\text{CNOT}}|\overline{0}\rangle_c \otimes |\overline{1}\rangle_t &= |\overline{0}\rangle_c \otimes |\overline{1}\rangle_t \\
\overline{\text{CNOT}}|\overline{1}\rangle_c \otimes |\overline{0}\rangle_t &= |\overline{1}\rangle_c \otimes |\overline{1}\rangle_t \\
\overline{\text{CNOT}}|\overline{1}\rangle_c \otimes |\overline{1}\rangle_t &= |\overline{1}\rangle_c \otimes |\overline{0}\rangle_t
\end{aligned}$$

Now we define the set $\{\overline{H}, \overline{T}, \overline{\text{CNOT}}, \overline{I}, \text{1-meas}, \text{1-prep}\}$ to be the set of 1-Ga's, and we call each element in this set a 1-Ga. Observe that the choice of our 1-Ga's given

C is not unique. We say that $\overline{\text{CNOT}}$ acts two blocks of encoded qubits, and all the other 1-Ga's act on one block of encoded qubits.

We define a 1-Rec to be a 1-Ga preceded by a 1-EC in each of its blocks. This definition is different from the one given in [2, 1] where they define a 1-Rec to be a 1-Ga followed by a 1-EC in each of its blocks, but is of no important consequence. Our rationale for doing this is only for pedagogy. Now we are in a position to introduce the concept of **concatenation** with respect to a quantum circuit. The notion of **concatenation** is important because it tells us explicitly how to construct a level-k circuit for $k \in \mathbb{Z}^+$.

Definition 2 (Concatenation) *Suppose that we want to incorporate error correction capabilities into some noisy quantum circuit $\tilde{\mathcal{Q}}$. We call all locations in $\tilde{\mathcal{Q}}$ 0-Ga. A level-1 version of $\tilde{\mathcal{Q}}$ is a quantum circuit that replaces every 0-Ga in $\tilde{\mathcal{Q}}$ with its equivalent 1-Rec. A 1-Rec corresponding to some 0-Ga that is not a 0-prep is defined as the corresponding 1-Ga preceded immediately by a 1-EC on each of the encoded blocks that the 1-Ga is designated to act on. A 1-Rec corresponding to a 0-prep is a 1-prep followed by a 1-EC. A 1-EC is a gadget that implements quantum error correction. A level-k version of $\tilde{\mathcal{Q}}$ is a quantum circuit that is obtained by repeating this replacement rule k times. We call this process of carrying out the replacement rule concatenation.*

4.4 Fault Tolerant Gadgets

We say a 1-Ga is fault-tolerant if it satisfies a list of properties that we will elaborate on later. But first we need to define a s-filter for $s \in \mathbb{N}$ following the terminology of [2]. We will define a s-filter using an operational language. If the input to a s-filter is equivalent to the ideal data acted on by a quantum operation of distance at most s, then the output of the s-filter is its input. If that is not the case, then the output will be the null state, indicating that the quantum computation is now unreliable.

Now we need to define a fault with respect to a quantum circuit. We say that a location in a quantum circuit has a fault if it is not implemented ideally.

In the previous section we have defined a 1-EC, a 1-Ga and a 1-Rec. Now we will define what a fault-tolerant 1-Ec and fault tolerant 1-Ga are. We remind the reader that at this point, 'fault-tolerant' should only be taken to be an adjective describing 1-Ga's that satisfy certain properties that will be formally stated. In the definitions that follow, we assume that we are using a quantum error correction code

that encodes one qubit with distance d. Let $t = \lfloor \frac{d-1}{2} \rfloor$ be the number of errors the quantum error correction code can correct. For brevity, let a 1-Ga or 1-EC that contains no more than r faults be called r-good.

Definition 3 (Fault tolerant 1-EC) *For all non-negative integers r', r, s such that $r' \leq t, r + s \leq t$, the following two properties are satisfied.*

1. –(r'-good 1-EC)– = –(r'-good 1-EC)–(r'-filter)–
2. –(s-filter)–(r-good 1-EC)–(i-decoder)– = –(s-filter)–(i-decoder)–

Note that the above definition is made using quantum circuit diagrams, where time is understood to flow from left to right.

Definition 4 (Fault tolerant 1-prep) *For all non-negative integer r such that $r \leq t$, the following two properties are satisfied.*

1. (r-good 1-prep)– = –(r-good 1-prep)–(r-filter)–
2. (r-good 1-prep)–(i-decoder)– = (ideal 0-prep)–

Definition 5 (Fault tolerant 1-meas) *For all non-negative integers r, s such that $r + s \leq t$, the following property is satisfied.*

1. –(s-filter)–(r-good 1-meas) = (s-filter)–(i-decoder)–(ideal 0-meas)

Definition 6 (Fault tolerant $\overline{\text{CNOT}}$) *For all non-negative integers r, s_1, s_2 such that $q = r + s \leq t$, the following two properties are satisfied where 1-Ga $\in \{\overline{I}, \overline{H}, \overline{T}\}$.*

1. $$\begin{matrix}(s_1\text{-filter})\text{–}\\ \\ (s_2\text{-filter})\text{–}\end{matrix}\left(\begin{matrix}\\ r\text{-good } \overline{\text{CNOT}} \\ \\ \end{matrix}\right)\text{–} = \begin{matrix}(s_1\text{-filter})\text{–}\\ \\ (s_2\text{-filter})\text{–}\end{matrix}\left(\begin{matrix}\\ r\text{-good } \overline{\text{CNOT}} \\ \\ \end{matrix}\right)\begin{matrix}\text{–}(s\text{-filter})\text{–}\\ \\ \text{–}(s\text{-filter})\text{–}\end{matrix}$$
2. $$\begin{matrix}(s_1\text{-filter})\text{–}\\ \\ (s_2\text{-filter})\text{–}\end{matrix}\left(\begin{matrix}\\ r\text{-good } \overline{\text{CNOT}} \\ \\ \end{matrix}\right)\begin{matrix}\text{–}(\text{i-decoder})\text{–}\\ \\ \text{–}(\text{i-decoder})\text{–}\end{matrix} = \begin{matrix}(s_1\text{-filter})\text{–}(\text{ i-decoder })\text{–}\\ \\ (s_2\text{-filter})\text{–}(\text{ i-decoder })\text{–}\end{matrix}\left(\begin{matrix}\\ \text{ideal } CNOT \\ \\ \end{matrix}\right)\text{–}$$

Definition 7 (Fault tolerant $\overline{I}$, $\overline{H}$, $\overline{T}$) *For all non-negative integers r, s such that $q = r + s \leq t$, the following two properties are satisfied where 1-Ga $\in \{\overline{I}, \overline{H}, \overline{T}\}$.*

1. –(s-filter)–(r-good 1-Ga)– = –(s-filter)–(r-good 1-Ga)–(q-filter)–
2. –(s-filter)–(r-good 1-Ga)–(i-decoder)– = –(s-filter)–(i-decoder)–(ideal 0-Ga)–

The intuition behind these definitions pertaining to fault-tolerance will become clearer in the next chapter. In particular, their definition is crucial to allow the **level reduction** as defined in the next chapter to go through.

Chapter 5

Local Noise, Level Reduction and Fault Tolerance

Under what conditions is it possible to implement a given quantum algorithm accurately using noisy gates? The goal of this chapter is to give sufficient conditions under which this is possible for the circuit model of quantum computation. Although there are many different sets of sufficient conditions for which this is possible, we will only restrict our attention to the one given by [2]. All of the results in this chapter come from [2, 1], and we merely give their results and proofs for pedagaogy.

5.1 Accuracy of Quantum algorithm

We wish to implement $\mathcal{Q}$ in the lab and call our noisy quantum algorithm $\tilde{\mathcal{Q}}$. We define our **bath** with Hilbert space $\mathcal{B}$ to be some set of particles not in our **system** such that the extension of all quantum operations acting on $\mathcal{S}$ are unitary in $\mathcal{S} \otimes \mathcal{B}$. Then $\tilde{\mathcal{Q}}$ is effectively the implementation of $\tilde{U} \in L(\mathcal{S} \otimes \mathcal{B})$ with initial state $|\psi\rangle = |0\rangle^{\otimes n}|\text{bath}\rangle$ and perfect measurement of $\tilde{U}|\psi\rangle$ in the computation basis.

Now let us denote the **accuracy** of a quantum algorithm $\tilde{\mathcal{Q}}$ with respect to an ideal algorithm $\mathcal{Q}$ as $Acc(\tilde{\mathcal{Q}}, \mathcal{Q})$. Let $\tilde{q}$ denote the probability distribution of the output of $\tilde{\mathcal{Q}}$. Then we define $Acc(\tilde{\mathcal{Q}}, \mathcal{Q}) = 1 - \|\tilde{q} - q\|_1$. It follows that $1 - Acc(\tilde{\mathcal{Q}}, \mathcal{Q}) \leq \|U \otimes V - (\tilde{U})\|$ for all unitary $V \in L(\mathcal{B})$ where $\|\cdot\|$ is the operator norm.

We can always write $\tilde{U}$ as $\tilde{U} = \prod_{\ell=1}^{L} U_\ell \Delta_\ell = \prod_{\ell=1}^{L} U_\ell (I_{\mathcal{S}} \otimes V_\ell + F_\ell)$ where Δ_ℓ is a unitary operator in $L(\mathcal{S} \otimes \mathcal{B})$, $V_\ell \in L(\mathcal{B}), F_\ell \in L(\mathcal{S} \otimes \mathcal{B})$ for all $\ell \in \{1, ..., L\}$ by use of

Lemma 1. The expansion of $\tilde{U}$ as a sum of 2^L terms in this way is called a **fault path expansion**, and each term in this expansion is called a **fault path**. We define the number of faults in a fault path to be the number of F_ℓ's in it. Many different fault path expansions can correspond to the same $\tilde{U}$. We make the important observation that any component of $U \otimes V - \tilde{U}$ must have at least one fault. Thus

$$\|U \otimes V - \tilde{U}\| \leq \sum_{k=1}^{L} \left\| \left(\prod_{j=1}^{k-1} U_j (I_{\mathcal{S}} \otimes V_j) \right) F_k \left(\prod_{\ell=k}^{L} U_k \Delta_\ell \right) \right\| = \sum_{\ell=1}^{L} \|F_\ell\| \qquad (5.1.1)$$

We will reproduce here the seminal result of [1, 2] that illustrates the crucial role the notion of a *local noise model* in the proof of the threshold theorem for fault tolerant quantum computing.

5.2 Local Noise

The follow definition is exactly what [2] uses and uses the same notation as far as possible.

Definition 8 (Local Noise) *Suppose that we have a system with Hilbert space $\mathcal{S}$ and a bath with Hilbert space $\mathcal{B}$. Suppose that we wish to simulate an ideal quantum algorithm $\mathcal{Q}$ with L locations which implements the unitary $U = \prod_{\ell=1}^{L} U_\ell \in L(\mathcal{S})$ using a noisy quantum circuit $\tilde{\mathcal{Q}}$. We let $\tilde{U} \in L(\mathcal{S} \otimes \mathcal{B})$ be the unitary transform that $\tilde{\mathcal{Q}}$ implements. Then we say that $\tilde{\mathcal{Q}}$ is subject to* **local noise** *with strength ϵ if there exists some fault path expansion for $\tilde{U}$ such that for all $r \in \{1, ..., L\}$ we have for all r-sets of locations $\mathcal{I}_r \subset \{1, ..., L\}$, $\|F(\mathcal{I}_r)\| \leq \epsilon^r$, where $F(\mathcal{I}_r)$ denotes the sum of all fault paths that have faults in all the locations indexed by $\mathcal{I}_r$.*

The following lemma has also been shown in [2], and we show it again for pedagogy.

Lemma 3 (Implication of local noise) *Consider a noisy quantum circuit $\tilde{\mathcal{Q}}$ with L locations with local noise strength ϵ. Let $\mathcal{L} = \{1, ..., n\}$ denote the set of all locations. Consider some $\mathcal{I}_C \subset \mathcal{L}$ where $|\mathcal{I}_C| = C$. Let*

$$S_{\tilde{\mathcal{Q}}} := \|\text{sum of all fault paths with at least } s \text{ faults in } \mathcal{I}_C\|$$

where the fault path expansion we choose is any one that allows us to show that $\tilde{\mathcal{Q}}$ has local noise with strength ϵ. Then $S_{\tilde{\mathcal{Q}}} \leq \binom{C}{s} \epsilon^s$.

Proof: We use the definition of local noise and the triangle inequality for norms.

Consider some $\mathcal{I} \subset \mathcal{I}_C$ where $|\mathcal{I}| = s$. By definition $F(\mathcal{I})$ is the sum of all fault paths with faults in all of the locations indexed by $\mathcal{I}$. This means that each fault path in $F(\mathcal{I})$ can also have faults in $\mathcal{L}\backslash\mathcal{I}$, which means that each fault path in $F(\mathcal{I})$ has at least s faults. This implies that $S_{\tilde{\mathcal{Q}}} = \| \sum_{\mathcal{I}\subset\mathcal{I}_C:|\mathcal{I}|=s} F(\mathcal{I})\|$. Since there are $\binom{C}{s}$ ways to pick s elements from the set $\mathcal{I}_\mathcal{C}$, and since the triangle inequality holds for any norm, and using the fact that the norm is always non-negative, we have $S_{\tilde{\mathcal{Q}}} \leq \binom{C}{s} \max_{\mathcal{I}\subset\mathcal{I}_C:|\mathcal{I}|=s}\{\|F(\mathcal{I})\|\}$. Since $\tilde{\mathcal{Q}}$ is subject to local noise with strength ϵ on L locations indexed by $\mathcal{L}$, it follows from definition of local noise that $\|F(\mathcal{I})\| \leq \epsilon^s$ for all $\mathcal{I}$ of order s. Thus $S_{\tilde{\mathcal{Q}}} \leq \binom{C}{s}\epsilon^s$. ∎

5.3 Level Reduction

Level reduction as introduced by [2, 1] in the context of fault-tolerant quantum computation is an important concept. The notion of **level reduction** is a mathematical tool that allows us to phrase the problem of decoding a level-$(k+1)$ circuit in terms of the properties of an equivalent level-k circuit. We reproduce the definition of level reduction as defined in [1]. In fact, all of the later terminology and proofs have been introduced by [2, 1].

Definition 9 (Level Reduction) *Suppose that we have a level-$(k+1)$ circuit subject to some noise model $\mathcal{N}^{(k+1)}$. We define the formal procedure of reducing this level-$(k+1)$ circuit with noise model $\mathcal{N}^{(k+1)}$ to a level-k circuit with noise model $\mathcal{N}^{(k)}$ as level reduction.*

Let an s-filter be a device that projects our state to one within s Pauli errors of a valid codeword. Let an i-decoder for a t-error correcting code be a device that corrects up to t errors and decodes the state it acts on to an unencoded qubit. We call a 1-Ga or 1-EC with at most r faults r-good.

Now we say a 1-Rec is **correct** if –(1-Rec)–(i-decoder)– = –(i-decoder)–(ideal 0-Ga)–.

Lemma 4 (Correctness) *Suppose that our 1-exRec is constructed using fault-tolerant gadgets, and has no more than t errors. Then the 1-Rec inside is correct.*

Proof: We only will give the proof in the case where the 1-Ga belongs to the set $\{\overline{I}, \overline{H}, \overline{T}\}$ because the proof for the other 1-Ga's is extremely similar. Consider non-negative integers r, s, s' such that $r + s + s' \leq t$. Then

–(s'-filter)–(s-good 1-EC)–(r-good 1-Ga)–(i-decoder)–
= –(s'-filter)–(s-good 1-EC)–(s-filter)–(r-good 1-Ga)–(i-decoder)–
= –(s'-filter)–(s-good 1-EC)–(s-filter)–(i-decoder)–(ideal 0-Ga)–
= –(s'-filter)–(s-good 1-EC)–(i-decoder)–(ideal 0-Ga)–
= –(i-decoder)–(ideal 0-Ga)–
= –(correct 1-Rec)– ∎

5.4 Invariance of Local Noise under Level Reduction

Lemma 5 (Local noise is preserved under level reduction) *Consider a noisy quantum circuit $\tilde{\mathcal{Q}}$ with L locations with local noise strength ϵ. Let $\mathcal{L} = \{1, ..., n\}$ denote the set of all locations. Let $\mathcal{C}$ be a partition of $\mathcal{L}$ with $c = |\mathcal{C}|$. Suppose that $\min_{c \in \mathcal{C}} |c| \geq 2(t+1)$. Then the operator norm of the sum of all fault paths that have at least $t+1$ faults in at least r elements of $\mathcal{C}$ is at most $\binom{C}{t+1}^r \left(\epsilon^{(t+1)}\right)^r$*

Proof: If $\min_{c \in \mathcal{C}} |c| \geq 2(t+1)$, then $\binom{C}{t+1} \geq \binom{|c|}{t+1} \forall c \in \mathcal{C}$ since $\binom{n}{t+1}$ increases monotonically with n for all $n \geq 2(t+1)$. Since $\mathcal{C}$ partitions $\mathcal{L}$, for any fault path, the occurence of at least $t+1$ faults in distinct partitions of $\mathcal{C}$ is independent. If a fault path has at least $t+1$ faults in each of r partitions of $\mathcal{C}$, then the fault path will have at least $r(t+1)$ faults. Then we apply Lemma 3 to obtain the result. ∎

Corollary 1 *A noisy quantum circuit $\tilde{\mathcal{Q}}$ of level $(k+1)$ with local noise strength ϵ is equivalent to a level (k) version with local noise strength $\epsilon' = \binom{C}{t+1}\epsilon^{(t+1)}$ where C is the maximal size of our 1-exRec and t is the number of errors that our quantum error correction code can correct.*

Proof: First we note that by the singleton bound, the number of qubits in each level-1 codeblock is necessarily greater than $4t$. This implies that $C > 2 \times 4t \geq (2t+1)$

for all $t \geq 1$ because the maximal size of a 1-exRec is greater than the number of qubits in two level-1 codeblocks. Hence we can use Lemma (4) in Lemma (5) to get the desired result. ■

5.5 Accuracy Threshold for Local Noise

Theorem 2 (Accuracy threshold for local noise) *Suppose that the 0-Ga's in our lab are subject to local noise with strength ϵ and that we wish to make a fault tolerant simulation of the ideal quantum algorithm $\mathcal{Q}$ with accuracy at least $1-\delta$. Assume that $\mathcal{Q}$ with L gates is constructed from gates that have fault-tolerant constructions with respect to some t-error correction code we choose. Then provided that $\epsilon < \epsilon_{thres} \leq \binom{C}{t+1}^{-1/t}$, this is possible by using k levels of concatenation where $k \geq \frac{\delta}{L\epsilon}\log(\epsilon/\epsilon_{thres})$. C is the maximum size of the 1-exRec we use.* ■

Proof: We have seen that Lemma 3 and 5 imply Corollary 1 which gives us a bound on ϵ_{thres}. Using the definition of accuracy with equation (5.1.1), we get a bound on k.

Now we know that any local noise model will give us a fault tolerant threshold by the above theorem. We can first write down the fault path expansion of quantum circuit at the unencoded level, and require that the supremum norm of any fault path with at least k faults is bounded from above by η^k where $\eta > 0$. We will then be able to apply the above theorem directly if $\eta < \epsilon_{thres}$. The remaining question is then, under what circumstances does such an η exist. Fortunately Lemma 2 in Chapter 3 already gives us an explicit upper bound for η. If the underlying non-Markovian noise model is also local, then η by definition is upper bounded by a constant that does not depend on how many qubits we use for our quantum error correction. In this case, we can directly apply the above theorem to get a threshold theorem for local non-Markovian noise.

Also we like to point out that in the proof of this threshold theorem, many implicit assumptions have been made. We will now make these assumptions explicit. We have assumed that quantum gates can be performed in parallel, and that there is either an inexhaustible supply of fresh ancilla qubits or the ability to refresh and reuse the ancilla qubits an indefinite number of times. We also assumed that CNOT can be performed between an arbitrary pair of qubits (even if they are not physically close together).

Chapter 6

Sharp Measurement

6.1 Motivation

We first define some terminology to make subsequent discussion less verbose. A system-bath Hamiltonian is defined as a Hermitian operator that acts non-trivially on what has been defined as the system and the bath. In this chapter we are interested in the specific case where the norm of the system-bath Hamiltonian due to measurement is parametrized by a positive real number g. If g is large, we will say that we have sharp measurement. We will define sharp measurements more precisely in the next section.

We have seen from the previous chapter that it is possible to perform fault-tolerant quantum computation given a local non-Markovian noise model if the amplitude for the local terms of the system-bath Hamiltonian are sufficiently small. For pedagogy, consider the following Hamiltonian model for quantum computation. Suppose every location in the quantum computer is implemented perfectly. In particular, we consider the case where the perfect measurements of qubits involve them being coupled strongly to the bath. Suppose that the the norm of the system-bath Hamiltonians associated with each of these couplings is very large. Then a direct upper bound on the noise strength of the local non-Markovian noise model will be very large, because any system-bath Hamiltonian is considered in this framework to be a contribution to noise. Hence in this special case, the noise strength of the local non-Markovian noise model will be too large for us to have a fault-tolerant threshold theorem for quantum computation.

An immediate objection to this example is that the system-bath interaction which

arises during the measurement process should not be considered as a noise process, but as an ideal operation in quantum computation. Thus a simple redefinition of which parts of the Hilbert space belong to the quantum computer, and which parts belong to the environment should easily yield a non-Markovian noise model that has zero noise strength for this particular example. Indeed, the objective of this chapter is to make this type of argument rigorous. In particular, we will need to make additional assumptions on the nature of the system-bath interaction during the measurement process to arrive at our result.

6.2 Redefinition of Noise Model

Now we proceed to do the most straightforward redefinition of the noise model. Let the Hilbert space of the universe be $\mathcal{H}_{\text{universe}}$. We divide the Hilbert space of the universe into four parts, which we label as $\mathcal{Q}, \mathcal{M}, \mathcal{B}$ and $\mathcal{C}$, so that $\mathcal{H}_{\text{universe}} = \mathcal{Q} \otimes \mathcal{M} \otimes \mathcal{C} \otimes \mathcal{B}$. We define the Hilbert space of our quantum computer to be $\mathcal{Q} \otimes \mathcal{C}$, and define the Hilbert space of our environment to be $\mathcal{B} \otimes \mathcal{M}$. $\mathcal{Q}$ is the Hilbert space of the part of the quantum computer. $\mathcal{C}$ is the Hilbert space of the classical registers that stores the result of measurements. $\mathcal{M}$ is the Hilbert space of the bath that is engineered specifically for purpose of measurement of qubits in $\mathcal{Q}$ and storing the measured data in $\mathcal{C}$. $\mathcal{B}$ is the bath for rest of the universe.

Now we describe how we model the measurement process. When a qubit in $\mathcal{Q}$ is measured, we allow the qubit to interact with a classical register in $\mathcal{C}$ and allow the measurement bath to interact with the measurement bath $\mathcal{M}$. We let the ideal system-bath interactions associated with this measurement process be H_{QC} and H_{CM} respectively. We let the non-ideal Hamiltonians associated with the measurement process to be H'_{QC} and H'_{CM} respectively.

Now assume that the Hamiltonian of the universe can be written as a sum of Hamiltonians for two-body interactions so that

$$H_{\text{universe}} = H_{\text{ideal}} + H_{\text{nonideal}} \tag{6.2.1}$$

where

$$H_{\text{ideal}} = H_Q + H_C + H_B + H_M + H_{QC} + H_{CM}$$

and

$$H_{\text{nonideal}} = H'_{QC} + H'_{CM} + H_{QB} + H_{QM} + H_{MB} + H_{CB}$$

where the subscript of each term in the equations above indicate the Hilbert spaces where the term concerned acts non-trivially. We assume that H_{ideal} is the part of the Hamiltonian that we want to implement and H_{nonideal} is the part of the Hamiltonian that we want to prevent. We define a noise Hamiltonian to be any term of H_{nonideal}.

Now we assume that the noise and ideal Hamiltonians are local in the sense that as the number of qubits in $\mathcal{Q}$ scales up, the norm of the fault of each location in $\mathcal{Q}$ can be upper bounded by an $\eta \in (0, 1)$. Let ℓ denote a location in $\mathcal{Q}$ and let $\mathcal{L}$ denote the set of all locations in $\mathcal{Q}$. We can write every noise Hamiltonian H' as $H' = \sum_{\sigma \in \mathcal{P}(\mathcal{L})} H_\sigma$ where $\mathcal{P}(\mathcal{L})$ is the power set of $\mathcal{L}$ and H_σ acts nontrivially on the set of location σ. For all $\ell \in \mathcal{L}$, define

$$H'_\ell := \sum_{\sigma \in \mathcal{P}(\mathcal{L}), \ell \in \sigma} H_\sigma.$$

Now suppose that for all $\ell \in \mathcal{L}$ and for all noise Hamiltonians H', $\|H'_\ell\|$ can be upper bounded by η'_H. We say equivalently that H' is local with strength η'_H. Let us suppose that all of the noise Hamiltonians are local.

Suppose each measurement takes time T, and all other location processes each take time $t_0 < \infty$. Without loss of generality we can have $t_0 \geq T$ for otherwise, we can just make the other locations wait for the measurement to be completed. We define H_{QC} and H_{CM} to be zero for time $t > T$ for each timestep, where t denotes the duration of the current timestep. Because of this fact, we will redefine $\eta_{H'_{QC}}$. In particular let $\eta_{H'_{QC,1}} = \sup_{t \in [0,T]} \|H'_{QC}(t)\|$ and $\eta_{H'_{QC,2}} = \sup_{t \in [T,t_0]} \|H'_{QC}(t)\|$ for all $\ell \in \mathcal{L}$.

We reiterate that we have assumed that the measurement process that arises from the H_{QC} and H_{CM} is perfect. The imperfections in the measurement process are quantified by the interaction terms H'_{QC} and H'_{CM}. Assume that the Hamiltonians for the ideal measurement process H_{QC} amd H_{CM} are local with strength that scales linearly with g. Since an ideal measurement process should entangle the qubit being measured to the measurement bath $\mathcal{M}$ and the classical register $\mathcal{C}$, we also assume that $gT = k$, where k is of order 1. Since we assume $\eta_{H'_{QC,1}}$ and $\eta_{H'_{CM}}$ scale linearly with g, it is reasonable to assume that $\eta_{H'_{QC,1}}$ and $\eta_{H'_{CM}}$ also scale linearly with g but such that $\eta_{H'_{QC,1}}/\eta_{H_{QC}}$ and $\eta_{H'_{CM}}/\eta_{H_{CM}}$ can be upper bounded by some small nonnegative constant δ. We assume that for all the other noise Hamiltonians H', the $\eta_{H'}$'s are independent of g.

6.3 The Result

By analyzing each timestep for all $\ell \in \mathcal{L}$, the Trotter decomposition, and using Lemma 2 we have ,

$$\begin{aligned}\|F_\ell\| &\leq (\eta_{H'_{QC,1}} + \eta_{H'_{CM}} + \eta_{H_{QB}} + \eta_{H_{QM}} + \eta_{H_{BM}})T/\hbar \\ &\quad + (\eta_{H_{QB}} + \eta_{H'_{QC,2}} + \eta_{H_{QM}})(t_0 - T)/\hbar \\ &= (\eta_{H'_{CM}} + \eta_{H'_{QC,1}} + \eta_{H_{BM}})T/\hbar + (\eta_{H_{QB}} + \eta_{H'_{QC,2}} + \eta_{H_{QM}})t_0/\hbar \qquad (6.3.1)\end{aligned}$$

Now using the fact that $gT = k$ we have for all $\ell \in \mathcal{L}$

$$\begin{aligned}\|F_\ell\| &\leq [(\eta_{H'_{CM}} + \eta_{H'_{QC,1}} + \eta_{H_{BM}})k/g + (\eta_{H_{QB}} + \eta_{H'_{QC,2}} + \eta_{H_{QM}})t_0]/\hbar \\ &= [2\delta k + \eta_{H_{BM}}k/g + (\eta_{H_{QB}} + \eta_{H'_{QC,2}} + \eta_{H_{QM}})t_0]/\hbar := \eta^*. \qquad (6.3.2)\end{aligned}$$

From the result that we have in Chapter 3, we can reduce our non-Markovian noise model to local noise model with strength at most η^*. From the above inequality, a large g (i.e. a sharp measurement) will cause the effective noise contribution from the H_{BM} interaction to be negligible. Our evaluation of η^* is the main result of this thesis. We emphasize that this result has been obtained by taking into account the dynamics of the measurement process explicitly and gives $\eta^* = 0$ for the 'pathological' example that we gave in the beginning of this chapter to motivate our work.

6.4 Conclusion

Although the results are not surprising, we emphasize that our analysis is the first that is done explicitly. It shows us the conditions under which we can still have a fault-tolerant threshold for quantum computation in the case where we explicitly take into account the problem of measurement. In the work of [1, 2], this was not done explicitly. However the nature of the bound that we have on the strength of our non-Markovian noise model is not fundamentally different from the result of [1, 2]. Recall that the two examples that we studied in Chapter 3 show that the size of the fault need not depend only on the size of the system-bath interaction. We hope to extend this intuition by coming up with specific phemonological noise models where the size of faults can become suppressed as some parameter in the noise model becomes large, such as the frequency of the perturbation or the ratio of the norms of the ideal Hamiltonian and the perturbation. We believe that this is an interesting open problem.

Appendices

The purpose of this appendix is to show the proof of Theorem 1. In the section on Fourier transforms and Bessel functions, we show how certain integrals are proportional to spherical Bessel functions. In the section on generating functions, we show how certain infinite summations are generating functions. In the section on binomial coefficients, we evaluate binomial identities that we will later use in Theorem 1. The section on the proof of Theorem 1 proves the theorem using material from the preceding three sections.

After the bulk of this thesis was written, we realized that there is a simpler proof to Theorem 1, and was also pointed out by one of our readers [36]. We will sketch the simple in the remaining part of this paragraph. Recall that $H, A \in L(\mathcal{H})$ are time-independent such that $HA = -AH$ and we wish to evaluate the propagator $e^{-i(H+A)t}$. The idea is to expand the propagator as a Taylor series. Since H and A anticommute, $(H + A)^2 = H^2 + A^2$. Also, it is esay to observe that $[H^2, A] = [H, A^2] = 0$. Now we separate out the even and the odd terms of the Taylor series expansion of our propagator, and factor an $(H+A)$ out of the each odd part. This factoring is possible since it involves only commuting terms. Hence we can obtain exactly the cosine and sinc terms needed in Theorem 1.

Although our original proof of this theorem is much longer, we believe that the techniques involved can be used to generalize our theorem to the case where anticommuting noise is time-dependent. Thus we still present our original proof of Theorem 1 here.

A Fourier Transforms and Bessel Functions

The following lemma is a trivial consequence of the definition of spherical Bessel functions and the relation between the Fourier transform of a particular function with Bessel functions [37].

Lemma 6 *For non-negative integer n,*

$$\int_{-1}^{1}(1-x^2)^n e^{isx}dx = j_n(s)\frac{2(n!)}{(s/2)^n}$$

where j_n is the order n spherical Bessel function of the first kind.

Proof: By definition from equation (10.1.1) in [38], we have

$$j_n(s) = \sqrt{\frac{\pi}{2s}}J_{n+\frac{1}{2}}(s). \tag{A.1}$$

where $J_{n+1/2}(s)$ is the Bessel function of the first kind of order $n+\frac{1}{2}$. Now from Appendix A of [37] we know that for $n \geq 0$,

$$\begin{aligned}\int_{-1}^{1}(1-x^2)^n e^{isx}dx &= J_{n+\frac{1}{2}}(s)\frac{n!\sqrt{\pi}}{(s/2)^{n+\frac{1}{2}}} \\ &= j_n(s)\sqrt{\frac{2s}{\pi}}\frac{n!\sqrt{\pi}}{(s/2)^{n+\frac{1}{2}}} \\ &= j_n(s)\sqrt{2s}\frac{n!}{(s/2)^n\sqrt{s/2}} \\ &= j_n(s)\frac{2(n!)}{(s/2)^n}. \quad \blacksquare\end{aligned} \tag{A.2}$$

We have not managed to find a proof of the following lemma in existing literature, but it is a trivial consequence of Lemma 6 and integration by parts.

Lemma 7 *For non-negative integer n,*

$$\int_{-1}^{1}x(1-x^2)^n e^{isx}dx = j_{n+1}(s)\frac{2in!}{(s/2)^n}$$

where j_n is the order n spherical Bessel function of the first kind.

Proof:

$$\begin{aligned}\int_{-1}^{1} x(1-x^2)^n e^{isx}dx &= \frac{(1-x^2)^{n+1}}{-2(n+1)}e^{isx}\Big|_{x=-1}^{x=1} - is\int_{-1}^{1}\frac{(1-x^2)^{n+1}}{-2(n+1)}e^{isx}dx \\ &= 0 + \frac{is}{2(n+1)}\int_{-1}^{1}(1-x^2)^{n+1}e^{isx}dx \\ &= \frac{is}{2(n+1)}j_{n+1}(s)\frac{2(n+1)!}{(s/2)^{n+1}} \\ &= j_{n+1}(s)\frac{2in!}{(s/2)^n} \quad \blacksquare\end{aligned}$$

B Generating Functions

From equation (10.1.40) of [38], we have a generating function for the spherical bessel functions.

$$\frac{1}{z}\cos\sqrt{z^2-2zt} = \sum_{n=0}^{\infty}\frac{t^n}{n!}j_{n-1}(z) \tag{B.1}$$

It follows that

$$\begin{aligned}\sum_{n=0}^{\infty}\frac{t^n}{n!}j_n(z) =& \frac{\partial}{\partial t}\frac{1}{z}\cos\sqrt{z^2-2zt} \\ =& \frac{1}{2}\cdot(-1)\cdot(-2z)\cdot\frac{1}{z}\cdot\frac{\sin\sqrt{z^2-2zt}}{\sqrt{z^2-2zt}} \\ =& \frac{\sin\sqrt{z^2-2zt}}{\sqrt{z^2-2zt}} := \operatorname{sinc}\sqrt{z^2-2zt}\end{aligned} \tag{B.2}$$

C Binomial Coefficients

Lemma 8 *Consider $k \in \mathbb{Z}^+$. Then*

$$\lim_{F\to\infty} \frac{1}{F^k}\binom{F}{k} = \frac{1}{k!}$$

Proof: Here k is constant, and we take the limit as F goes to ∞. By definition,

$$\binom{F}{k} = \frac{F_{(k)}}{k!}$$

where we are using the Pochhammer symbol in the numerator of the expression on the left hand side of the above equation. Since k is constant,

$$\lim_{F\to\infty} \frac{1}{F^k}\binom{F}{k} = \lim_{F\to\infty} \frac{1}{F^k}\frac{F_{(k)}}{k!} = \frac{1}{k!}\lim_{F\to\infty}\prod_{j=0}^{k-1}\left(1-\frac{j}{F}\right) = \frac{1}{k!} \quad \blacksquare$$

Lemma 9 *Consider $k = \kappa F \in \mathbb{Z}^+$ where $\kappa \in (0,1), F \in \mathbb{Z}^+$. Then*

$$\lim_{F\to\infty} \frac{1}{F^k}\binom{F}{k} = 0$$

Proof: The Stirling approximation is

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}(n/e)^n} = 1$$

Since F, k and $F-k$ all approach infinity as F approaches infinity, we can apply the Stirling approximation to $\binom{F}{k} = \frac{F!}{k!(F-k)!}$ to get

$$\begin{aligned}
\lim_{F\to\infty} \frac{1}{F^k}\binom{F}{k} &= \lim_{F\to\infty} \frac{1}{F^k}\frac{F!}{k!(F-k)!} = \lim_{F\to\infty} \frac{1}{F^k}\frac{F^F}{k^k(F-k)^{F-k}}\sqrt{\frac{F}{2\pi k(F-k)}} \\
&= \lim_{F\to\infty} \frac{F^{F-k}}{k^k(F-k)^{F-k}}\sqrt{\frac{F}{2\pi k(F-k)}} \\
&= \lim_{F\to\infty} \kappa^{\kappa F}(1-\kappa)^{(1-\kappa)F}\sqrt{\frac{1}{2\pi\kappa(1-\kappa)F}} = 0.
\end{aligned}$$

The last equality holds because κ and $(1-\kappa)$ are real constants with absolute value strictly less than 1. $\blacksquare$

Lemma 10 *Consider $k = F - a \in \mathbb{Z}^+$ where $a \in \mathbb{Z}^+$. Then*

$$\lim_{N\to\infty} \frac{1}{F^k}\binom{F}{k} = 0$$

Proof: We again use the Stirling approximation for the factorial terms in the binomial coeffient that become arbitrarily large. Then we have

$$\lim_{N\to\infty} \frac{1}{F^k}\binom{F}{k} = \lim_{F\to\infty} \frac{1}{F^k}\frac{F!}{k!(F-k)!} = \lim_{F\to\infty} \frac{1}{F^k}\frac{F^F}{k^k a!}\sqrt{\frac{F}{k}} = \lim_{F\to\infty} \frac{1}{F^F}\frac{1}{a!} = 0 \quad \blacksquare$$

Lemma 11 *Let N, L be positive even integers and $-N \leq L \leq N$. Let $\ell = L/N$. Then (i) for $k \in \mathbb{Z}^+$ being constant*

$$\lim_{N\to\infty} N\binom{\frac{N+L}{2}-1}{k-1}\binom{\frac{N-L}{2}}{k-1}\left(\frac{1}{N}\right)^{2k-1} = \frac{\left(\frac{1-\ell^2}{4}\right)^{k-1}}{(k-1)!(k-1)!}$$

and (ii) for $k = \kappa N$ where $\kappa \in (0,1)$

$$\lim_{N\to\infty} N\binom{\frac{N+L}{2}-1}{k-1}\binom{\frac{N-L}{2}}{k-1}\left(\frac{1}{N}\right)^{2k-1} = 0.$$

Proof:

$$\begin{aligned}
&\lim_{N\to\infty} N\binom{\frac{N+L}{2}-1}{k-1}\binom{\frac{N-L}{2}}{k-1}\left(\frac{1}{N}\right)^{2k-1} \\
&= \lim_{N\to\infty} N\binom{F-1}{k-1}\binom{B}{k-1}\left(\frac{1}{N}\right)^{2k-1} \\
&= \lim_{N\to\infty} N\binom{F}{k-1}\binom{B}{k-1}\left(\frac{1}{N}\right)^{2k-1} \\
&= \lim_{N\to\infty} \frac{1}{N^{k-1}}\binom{F}{k-1}\frac{1}{N^{k-1}}\binom{B}{k-1} \\
&= \lim_{N\to\infty} \frac{((1+\ell)/2)^{k-1}}{F^{k-1}}\binom{F}{k-1}\frac{((1-\ell)/2)^{k-1}}{B^{k-1}}\binom{B}{k-1} \\
&= \left(\frac{1-\ell^2}{4}\right)^{k-1} \lim_{N\to\infty} \frac{1}{F^{k-1}}\binom{F}{k-1}\frac{1}{B^{k-1}}\binom{B}{k-1} \qquad \text{(C.1)}
\end{aligned}$$

If $k = 1$, then

$$\lim_{N\to\infty} \frac{1}{F^{k-1}}\binom{F}{k-1}\frac{1}{B^{k-1}}\binom{B}{k-1} = 1$$

If $k > 1$, first consider the case where $|L| = N - 2a$ where a is some positive constant integer. Then

$$\lim_{N\to\infty} \frac{1}{F^{k-1}}\binom{F}{k-1}\frac{1}{B^{k-1}}\binom{B}{k-1} = \frac{1}{a^{k-1}}\binom{a}{k-1}\lim_{N\to\infty}\frac{1}{N^{k-1}}\binom{N}{k-1} = 0.$$

In the case where $L = \ell N$, and k is constant, by using Lemma 8 we get

$$\lim_{N\to\infty} \frac{1}{F^{k-1}}\binom{F}{k-1}\frac{1}{B^{k-1}}\binom{B}{k-1} = \frac{1}{(k-1)!(k-1)!}$$

In the case where $L = \ell N$, and $k = \kappa N$ for $\kappa \in (0,1)$ being constant, by using Lemma 9, 10 we get

$$\lim_{N\to\infty} \frac{1}{F^{k-1}}\binom{F}{k-1}\frac{1}{B^{k-1}}\binom{B}{k-1} = 0.$$

These results combined with equation C.1 proves our lemma. ∎

Lemma 12 *Let N, L be positive even integers and $-N \le L \le N$. Let $\ell = L/N$. Then (i) for $k \in \mathbb{Z}^+$ being constant*

$$\lim_{N\to\infty} N\binom{\frac{N+L}{2}}{k}\binom{\frac{N-L}{2}}{k-1}\left(\frac{1}{N}\right)^{2k} = \left(\frac{1+\ell}{2}\right)\frac{\left(\frac{1-\ell^2}{4}\right)^{k-1}}{k!(k-1)!}$$

and (ii) for $k = \kappa N \in \mathbb{Z}^+$ where $\kappa \in (0,1)$

$$\lim_{N\to\infty} N\binom{\frac{N+L}{2}}{k}\binom{\frac{N-L}{2}}{k-1}\left(\frac{1}{N}\right)^{2k} = 0$$

Proof: Let $F = \frac{N+L}{2}$ and $B = \frac{N-L}{2}$. Then $F = (1+\ell)N/2, B = (1-\ell)N/2$. Thus

$$\begin{aligned}
&\lim_{N\to\infty} N\binom{\frac{N+L}{2}}{k}\binom{\frac{N-L}{2}-1}{k-1}\left(\frac{1}{N}\right)^{2k} \\
&= \lim_{N\to\infty} N\binom{F}{k}\binom{B-1}{k-1}\left(\frac{1}{N}\right)^{2k} \\
&= \lim_{N\to\infty} N\binom{F}{k}\binom{B}{k-1}\left(\frac{1}{N}\right)^{2k} \\
&= \lim_{N\to\infty} \frac{1}{N^k}\binom{F}{k}\frac{1}{N^{k-1}}\binom{B}{k-1} \\
&= \lim_{N\to\infty} \frac{((1+\ell)/2)^k}{F^k}\binom{F}{k}\frac{((1-\ell)/2)^{k-1}}{B^{k-1}}\binom{B}{k-1} \\
&= \left(\frac{1-\ell^2}{4}\right)^{k-1}\left(\frac{1+\ell}{2}\right)\lim_{N\to\infty}\frac{1}{F^k}\binom{F}{k}\frac{1}{B^{k-1}}\binom{B}{k-1}. \qquad \text{(C.2)}
\end{aligned}$$

First suppose that $|L| = N - 2a$ for some constant $a \in \mathbb{Z}^+$ so that as $N \to \infty$ we have $\ell \to -1$ or $\ell \to 1$. Then for $k = 1$,

$$\lim_{N\to\infty}\frac{1}{F^{k-1}}\binom{F}{k-1}\frac{1}{B^{k-1}}\binom{B}{k-1} = 1$$

and for integer $k > 1$,

$$\lim_{N\to\infty}\frac{1}{F^{k-1}}\binom{F}{k-1}\frac{1}{B^{k-1}}\binom{B}{k-1} = \frac{1}{a^{k-1}}\binom{a}{k-1}\lim_{N\to\infty}\frac{1}{N^{k-1}}\binom{N}{k-1} = 0.$$

Now we consider $\ell \in (0,1)$. In the case where k is a constant integer, by Lemma 8 we get

$$\lim_{N\to\infty}\frac{1}{F^k}\binom{F}{k}\frac{1}{B^{k-1}}\binom{B}{k-1} = \frac{1}{k!(k-1)!}.$$

In the case where $k = \kappa N$ for $\kappa \in (0,1)$, using Lemma $9, 10$ we get

$$\lim_{N\to\infty}\frac{1}{F^k}\binom{F}{k}\frac{1}{B^{k-1}}\binom{B}{k-1} = 0.$$

These results combined with equation C.2 proves our lemma. ∎

D Proof of Theorem 1

Let us give the statement of Theorem 1 again.

Theorem 1 *Suppose that we have a separable Hilbert space $\mathcal{H}$ and have Hermitian operator $H \in L(\mathcal{H})$ and a linear operator $A \in L(\mathcal{H})$ such that $HA = -AH$. Let $H = \sum_m h_m |m\rangle\langle m|$ be the spectral decomposition of H. Let $\mathbb{1}$ be the identity operator on $\mathcal{H}$. Then for $t > 0$,*

$$e^{-i(H+A)t} = \sum_m |m\rangle\langle m| \left(\cos(t\sqrt{h_m^2 \mathbb{1} + A^2}) - i(h_m \mathbb{1} + A)t \operatorname{sinc}(t\sqrt{h_m^2 \mathbb{1} + A^2}) \right) \tag{D.1}$$

where $\operatorname{sinc}(x) := \sin(x)/x$ *is a formal power series in* x.
Now let $H_a = H + A$, and $F_a = e^{-iH_a t} - e^{-iHt}$. Then $F_a = U_a - U$. Let $\| \cdot \|$ be the operator norm.

First we use the Trotter product formula to express our exponential as a sum of infinitely many monomials. We then count the number of our monomials that satisfy certain properties, and thereby express our original exponential as an integral with basis states being the eigenvectors of our unperturbed Hamiltonian. We then realize that the Fourier transforms that come up are directly proportional to the spherical Bessel functions. Thus we obtain a summation over the spherical Bessel functions. We then observe that the summations that we have at hand are precisely the generating functions of some special series. Thus our F_a is expressed eventually as linear combination of generating functions which are easy to analyze. After this overview, let us begin with the detailed version of the calculation.

By the Trotter product formula [30], we have

$$U_a = \lim_{N \to \infty} \prod_{j=1}^{N} e^{-iHt/N}(I - iAt/N) \tag{D.2}$$

Without loss of generality, we will assume that N is always an even integer. First we expand the n-fold product above. We will associate each monomial in the expansion with a diagram which we call a **KL-diagram** for lack of a better name. We also associate each KL-diagram with a KL which is just an ordered pair of integers that keeps track of the more important properties of a KL-diagram. A KL-diagram is just a pictoral way of visualizing each monomial. We will introduce some terms associated with a KL-diagram to make some of the calculations we do later more transparent.

We define a KL-diagram by an example. In the table below we give a few examples of monomials in the Trotter expansion of (D.2) with their corresponding KL-diagrams in the case where $N = 4$. Let $G = e^{-iHt/N}$.

monomial	KL-diagram
$GGGG$	`x-.-.-.-o`
$G(-iAt/N)GGG$	`x-A-.-.-.o`
$GG(-iAt/N)GG$	`x-.-A-.-.o`
$GGG(-iAt/N)G$	`x-.-.-A-.o`
$GGGG(-iAt/N)$	`x-.-.-.-Ao`
$G(-iAt/N)G(-iAt/N)GG$	`x-A-A-.-.o`

We now introduce all the terminology we need for our KL-diagram. The `'x'` in the KL-diagram is just indicates the beginning of the diagram. The `'-'` in the KL-diagram represents G. The `'.'` in the KL-diagram represents I. The `'A'` in the KL-diagram represents $-itA/N$. The `'o'` in the KL-diagram indicates the end of the diagram. We call the string of symbols between consecutive `A`s, between `'x'` and `'A'`, between `'A'` and `'o'` as **segments**. The first segment is the one that starts after `'x'` and the last segment is terminated by `'o'`. We will label every segment in a KL-diagram with a number, where the first segment has **segment number** equals to 1 and the last segment will have **segment number** equals to N. Let K be the number of `'A'`s in a KL-diagram. (This motivates the use of 'K' in KL-diagram). Then the number of segments in each KL-diagram is $K + 1$. If the a segment of a KL-diagram has an odd segment number, we say that it is a forward segment, otherwise it is a backward segment. The length of each segment is the number of dashes in the segment. Let the sum of the length of all the forward segments be F and the sum of all the backward segments be B. We define $L = F - B$. L is to be interpreted as the displacement of a squirrel from the origin if it hops F steps forwards and B steps backwards. Here, 'L' is used in the name of the KL-diagram because each such diagram has a L associated with it. We call the length of the first forward segment F_1, the length of the second forward segment F_2 and so on. We call the length of the first backward segment B_1, the length of the second backward segment B_2 and so on. We observe that every segment has a length of at least one except for the last segment. Let n_F be the number of forward segments and n_B be the number of backward segments. If K is even, then $n_F = K/2 + 1, n_B = K/2$ and the last segment is forwards. If K is odd, then $n_F = (K+1)/2, n_B = (K+1)/2$ and the last segment is backwards.

For each monomial in the Trotter expansion, we will move all the $e^{-iHt/N}$s to the left hand side by using the identity $Ae^{\pm iHt/N} = e^{\mp iHt/N}A$ repeatedly. We then

combine all the terms that are powers $e^{-iHt/N}$. For a monomial with number of '`A`'s being K and displacement L, this monomial is precisely $e^{-iHtL/N}(-itA/N)^K$. The question that we would like to answer is: given fixed K and L, how many such monomials do we have in our Trotter expansion? Let the answer to this combinatorial question be $c_{K,L}$.

Then

$$U_a = U + \lim_{N\to\infty} \sum_{K=1}^{N} \sum_{L \text{ even}} c_{K,L} e^{-iHtL/N} (-itA/N)^K \tag{D.3}$$

where $L \in [-N+2, N]$. Now let us evaluate $c_{K,L}$. Consider any KL-diagram with displacement L and K number of `A`s where K is even. Then the corresponding number of forward segments is $n_F = K/2 + 1$ and the number backwards segments is $n_B = K/2$. The last segment is forwards, and thus the last segment can have a length of zero. All other segments have length of at least one. In the case where K is odd, $n_F = n_B = (K+1)/2$ and the last segment is backwards. Observe that $F = (N+L)/2$ and $B = (N-L)/2$. Then from the combinatorial fact that there are $\binom{n-1}{k}$ ways to put k separators between n objects arranged in a line such that there is at least one object between each separator, $c_{K,L}$ is given by

$$c_{K,L} = \begin{cases} \binom{F-1}{n_F-1}\binom{B}{n_B-1} & K \text{ odd} \\ \binom{F}{n_F-1}\binom{B-1}{n_B-1} & K \text{ even} \end{cases} \tag{D.4}$$

Now we will consider four separate cases when $N \to \infty$. The four cases are

1. $\lim_{N\to\infty} K/N = 0$, $\lim_{N\to\infty} L/N = 0$
2. $\lim_{N\to\infty} K/N \neq 0$, $\lim_{N\to\infty} L/N = 0$
3. $\lim_{N\to\infty} K/N \neq 0$, $\lim_{N\to\infty} L/N \neq 0$
4. $\lim_{N\to\infty} K/N = 0$, $\lim_{N\to\infty} L/N \neq 0$

By use of Lemmas 11 and 12, only Case 4 has a non-zero contribution as $N \to \infty$.

Thus

$$
\begin{aligned}
F_a &= \lim_{N\to\infty} \sum_{K=1}^{N} \int_{L=-N}^{L=N} \frac{dL}{2} e^{-iHtL/N} c_{K,L} (-itA/N)^K \\
&= \lim_{N\to\infty} \sum_{K=1}^{N} \int_{\ell=-1}^{\ell=1} \frac{Nd\ell}{2} e^{-iHtL/N} c_{K,\ell N} (-itA/N)^K \\
&= \lim_{N\to\infty} \sum_{k=1}^{N/2} \int_{\ell=-1}^{\ell=1} \frac{Nd\ell}{2} e^{-iHtL/N} \left(\frac{c_{2k-1,\ell N}}{N^{2k-1}} (-itA)^{2k-1} + \frac{c_{2k,\ell N}}{N^{2k}} (-itA)^{2k} \right) \\
&= \frac{1}{2} \sum_{k=1}^{\infty} \int_{-1}^{1} d\ell e^{-iHt\ell} \left(\frac{\left(\frac{1-\ell^2}{4}\right)^{k-1} (-itA)^{2k-1}}{(k-1)!(k-1)!} + \frac{\left(\frac{1+\ell}{2}\right) \left(\frac{1-\ell^2}{4}\right)^{k-1} (-itA)^{2k}}{k!(k-1)!} \right) \\
&= \frac{1}{2} \sum_{k=1}^{\infty} \int_{-1}^{1} d\ell e^{-iHt\ell} \left(\frac{(1-\ell^2)^{k-1} \left(\frac{-itA}{2}\right)^{2k-1} \cdot 2}{(k-1)!(k-1)!} + \frac{(1+\ell)(1-\ell^2)^{k-1} \left(\frac{-itA}{2}\right)^{2k} \cdot 2}{k!(k-1)!} \right) \\
&= \sum_{k=1}^{\infty} \int_{-1}^{1} d\ell e^{-iHt\ell} \left(\frac{(1-\ell^2)^{k-1} \left(\frac{-itA}{2}\right)^{2k-1}}{(k-1)!(k-1)!} + \frac{(1+\ell)(1-\ell^2)^{k-1} \left(\frac{-itA}{2}\right)^{2k}}{k!(k-1)!} \right).
\end{aligned}
\tag{D.5}
$$

Recall that $H = \sum_m h_m |m\rangle\langle m|$ where $h_m \in \mathbb{R}$ and $\{|m\rangle\}$ is an orthonormal basis. Thus $e^{-iHt\ell} = \sum_m e^{-ih_m t\ell} |m\rangle\langle m|$. Also by the fact that (-1) raised to an even power is $+1$ and (-1) raised to an odd power is (-1) we get

$$
\begin{aligned}
F_a = &\sum_{k=1}^{\infty} \int_{-1}^{1} d\ell \sum_m e^{-ih_m t\ell} |m\rangle\langle m| \\
&\times \left(\frac{-(1-\ell^2)^{k-1}}{(k-1)!(k-1)!} \left(\frac{itA}{2}\right)^{2k-1} + (1+\ell) \frac{(1-\ell^2)^{k-1}}{k!(k-1)!} \left(\frac{itA}{2}\right)^{2k} \right).
\end{aligned}
$$

We first exchange the order of the summation and the integral. Then we get

$$
\begin{aligned}
F_a = &\sum_m |m\rangle\langle m| \sum_{k=1}^{\infty} \int_{-1}^{1} d\ell e^{-ih_m t\ell} \\
&\times \left(\frac{-(1-\ell^2)^{k-1}}{(k-1)!(k-1)!} \left(\frac{itA}{2}\right)^{2k-1} + (1+\ell) \frac{(1-\ell^2)^{k-1}}{k!(k-1)!} \left(\frac{itA}{2}\right)^{2k} \right).
\end{aligned}
$$

Using Lemma 6 and Lemma 7 we can evaluate some of the integrals of the above equation to spherical Bessel functions to get

$$F_a = \sum_m |m\rangle\langle m| \sum_{k=1}^{\infty} \left(\frac{-2j_{k-1}(-h_m t)}{(-h_m t/2)^{k-1}(k-1)!} \left(\frac{itA}{2}\right)^{2k-1} + \frac{2j_{k-1}(-h_m t) + 2ij_k(-h_m t)}{(-h_m t/2)^{k-1} k!} \left(\frac{itA}{2}\right)^{2k} \right). \tag{D.6}$$

Simplifying the above equation we get

$$\begin{aligned} F_a &= \sum_m |m\rangle\langle m| \sum_{k=1}^{\infty} \left(\frac{-2j_{k-1}(-h_m t)}{(-h_m t/2)^{k-1}(k-1)!} \left(\frac{itA}{2}\right)^{2k-2} \left(\frac{itA}{2}\right) + \frac{2j_{k-1}(-h_m t) + 2ij_k(-h_m t)}{(-h_m t/2)^{k-1} k!} \left(\frac{itA}{2}\right)^{2k} \right) \\ &= \sum_m |m\rangle\langle m| \sum_{k=1}^{\infty} \left(\frac{j_{k-1}(-h_m t)}{(k-1)!} \left(\frac{2}{-h_m t}\frac{-t^2A^2}{4}\right)^{k-1} \left(\frac{-2itA}{2}\right) + \frac{j_{k-1}(-h_m t) + ij_k(-h_m t)}{k!} \left(\frac{2}{-h_m t}\frac{-t^2A^2}{4}\right)^{k} \frac{-h_m t}{2} \cdot 2 \right) \\ &= \sum_m |m\rangle\langle m| \left(-itA \sum_{k=0}^{\infty} \frac{j_k(-h_m t)}{k!} \left(\frac{t^2A^2}{2h_m t}\right)^k - h_m t \sum_{k=1}^{\infty} \frac{j_{k-1}(-h_m t) + ij_k(-h_m t)}{k!} \left(\frac{t^2A^2}{2h_m t}\right)^k \right). \end{aligned} \tag{D.7}$$

Using the generating functions we have for the spherical Bessel functions (B.1), (B.2), we get

$$\begin{aligned} F_a = \sum_m |m\rangle\langle m| \Big(& -itA\,\mathrm{sinc}(t\sqrt{h_m^2 \mathbb{1} + A^2}) \\ & + \left(\cos(t\sqrt{h_m^2\mathbb{1} + A^2}) - \mathbb{1}\cos(h_m t)\right) \\ & - ih_m t \left(\mathrm{sinc}(t\sqrt{h_m^2\mathbb{1} + A^2}) - \mathbb{1}\,\mathrm{sinc}(h_m t)\right)\Big). \end{aligned} \tag{D.8}$$

Simplifying the above equation we get

$$\begin{aligned} e^{-iHt} + F_a &= \sum_m |m\rangle\langle m| \left(-i(A + h_m\mathbb{1})t\,\mathrm{sinc}(t\sqrt{h_m^2\mathbb{1} + A^2}) + \cos(t\sqrt{h_m^2\mathbb{1} + A^2}) \right) \\ e^{-i(H+A)t} &= \sum_m |m\rangle\langle m| \left(\cos(t\sqrt{h_m^2\mathbb{1} + A^2}) - i(h_m\mathbb{1} + A)t\,\mathrm{sinc}(t\sqrt{h_m^2\mathbb{1} + A^2}) \right) \end{aligned} \tag{D.9}$$

and this proves our theorem. ■

Bibliography

[1] P. Aliferis, D. Gottesman, and J. Preskill, "Quantum accuracy threshold for concatenated distance-3 codes," *Quant. Inf. Comput.*, vol. 6, pp. 97–165, 2006. quant-ph/0504218.

[2] P. Aliferis, *Level Reduction and the Quantum Threshold Theorem*. PhD thesis, California Institute of Technology, 2007. quant-ph/0703230.

[3] P. O. Boykin, T. Mor, M. Pulver, V. Roychowdhury, and F. Vatan, "On universal and fault-tolerant quantum computing," (Los Alamitos, CA), IEEE Computer Society Press, 1999. quant-ph/9906054v1.

[4] R. P. Feynman, "Simulating physics with computers," *Int. J. Theor. Phys.*, vol. 21, pp. 467–488, 1982.

[5] R. P. Feynman, "Quantum mechanical computers," *Foundations of physics*, vol. 16, pp. 507–531, 1986.

[6] P. Benioff, "Quantum mechanical Hamiltonian models of Turing machines," *J. Stat. Phys.*, vol. 29, pp. 515–546, 1982.

[7] D. Deutsch, "Quantum theory, the Church-Turing principle and the universal quantum computer," *Proc. R. Soc. London A*, vol. 400, pp. 97–117, 1985.

[8] D. S. Abrams and S. Lloyd, "Simulations of many-body Fermi systems on a universal quantum computer," *Phys. Rev. Lett.*, vol. 79, pp. 2586–2589, 1997. quant-ph/9703054.

[9] M. H. Freedman, A. Kitaev, and Z. Wang, "Simulation of topological field theories by quantum computers," *Commun. Math. Phys.*, vol. 227, pp. 587–603, 2002.

[10] M. Byrnes and Y. Yamamoto, "Simulating lattice gauge theories on a quantum computer," *Phys.Rev. A*, vol. 73, p. 022328, 2006. quant-ph/0510027v1.

[11] D. S. Abrams and S. Lloyd, "Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors," *Phys. Rev. Lett.*, vol. 83, pp. 5162–5165, 1999. quant-ph/9807070.

[12] T. Szkopek, V. Roychowdhury, E. Yablonovitch, and D. S. Abrams, "Eigenvalue estimation of differential operators," *Phys. Rev. A*, vol. 72, p. 062318, 2005. quant-ph/0408137v4.

[13] C. Zalka, "Simulating quantum systems on a quantum computer," *Proc R Soc London Ser A*, vol. 454, pp. 313–322, 1998.

[14] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik, "Polynomial-time quantum algorithm for the simulation of chemical dynamics," *Proc. of the Natural Academy of Sciences (PNAS)*, vol. 105, pp. 18681–18686, 2008.

[15] P. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM J.Sci.Statist.Comput.*, vol. 26, p. 1484, 1997.

[16] L. K. Grover, "A fast quantum mechanical algorithm for database search," *Proc. 28th Annual Symposium on the Theory of Computing*, pp. 212–218, 1996. quant-ph/9605043.

[17] R. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public key crytosystems," *Comm. of the ACM*, vol. 21, pp. 120–126, 1978.

[18] RSA, The Security Division of EMC, http://www.rsa.com/node.aspx?id=1274.

[19] C. Negrevergne, T. S. Mahesh, C. A. Ryan, M. Ditty, F. Cyr-Racine, W. Power, N. Boulant, T. Havel, D. G. Cory, and R. Laflamme, "Benchmarking quantum control methods on a 12-qubit system," *Phys. Rev. Lett.*, vol. 96, p. 170501, 2006.

[20] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood, and I. L. Chuang, "Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance," *Nature*, vol. 414, pp. 883–887, 2001.

[21] B. M. Terhal and G. Burkard, "Fault-tolerant quantum computation for local non-markovian noise," *Phys. Rev. A*, vol. 71, p. 012336, 2005. quant-ph/0402104.

[22] B. W. Reichardt, "Fault-tolerance threshold for a distance-three quantum code," 2005. quant-ph/0509203.

[23] P. Aliferis and J. Preskill, "Fault-tolerant quantum computation against biased noise," *Phys. Rev. A*, vol. 78, p. 052331, 2008. quant-ph/0710.1301.

[24] P. Aliferis and J. Preskill, "The fibonacci scheme for fault-tolerant quantum computation," 2008. quant-ph/0809.5063.

[25] D. Aharonov, A. Kitaev, and J. Preskill, "Fault-tolerant quantum computation with long-range correlated noise," *Phys. Rev. Lett.*, vol. 96, p. 050504, 2006. quant-ph/0510231.

[26] R. Raussendorf, J. Harrington, and K. Goyal, "A fault-tolerant one-way quantum computer," *Annals of Physics*, vol. 321, p. 2242, 2006. quant-ph/0510135.

[27] F. Schmidt-Kaler, H. Häffner, M. Riebe, S. Gulde, G. P. T. Lancaster, T. Deuschle, C. Becher, C. F. Roos, J. Eschner, and R. Blatt, "Realization of the Cirac-Zoller controlled-not quantum gate," *Nature*, vol. 422, pp. 408–411, 2002.

[28] A. J. Leggett, S. Chakravarty, A. T. Dorsey, M. P. A. Fisher, and W. Z. Anupam Garg, "Dynamics of the dissipative two-state system," *Reviews of Modern Physics*, vol. 59, no. 1, 1987. quant-ph/0610063.

[29] R. Shankar, *Principles of Quantum Mechanics*. 233 Spring Street, New York, N.Y. 100130: Plenum Press, second ed., 1994.

[30] M. Reed and B. Simon, *Methods of Mathematical Physics I: Functional Analysis*. New York and London: Academic Press, first ed., 1972.

[31] K. Kraus, *Lecture Notes in Physics 190 : States, Effects, and Operations Fundamental Notions of Quantum Theory*. Springer Berlin / Heidelberg, first ed., 1983.

[32] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, second ed., 2000.

[33] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*. Great Clarendon Street, Oxford OX2 6DP, UK: Oxford University Press, first ed., 2002.

[34] E. Knill and R. Laflamme, "A theory of quantum error correcting codes," *Phys. Rev. A*, vol. 55, p. 900, 1997. quant-ph/9604034.

[35] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed state entanglement and quantum error correction," *Phys. Rev. A*, vol. 54, pp. 3824–3851, 1996. quant-ph/9604024.

[36] A. M. Childs, private communication.

[37] E. M. Stein and R. Shakarchi, *Complex Analysis*. Princeton: Princeton University Press, first ed., 2003.

[38] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover, ninth dover printing, tenth gpo printing ed., 1964.